

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:13 ; Search time 244.117 Seconds
(without alignments)
10530.153 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571

Sequence: 1 eggsgaggtacgggggaattga.....cccagaataactgagag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1567.8	99.8	10818	4 US-09-949-016-13583	Sequence 13583, A
2	599	38.1	601	4 US-09-949-016-63416	Sequence 63416, A
3	482	30.7	735	4 US-09-602-877A-94	Sequence 94, Appl
4	482	30.7	1143	4 US-09-949-016-1841	Sequence 1841, Ap
5	482	30.7	1499	4 US-09-509-908-1	Sequence 1, Appl1
6	480.6	30.6	601	4 US-09-949-016-63415	Sequence 63415, A
7	476	30.3	1476	2 US-08-824-874-2	Sequence 2, Appl1
8	476	30.3	1476	2 US-09-210-084-2	Sequence 2, Appl1
9	476	30.3	1476	4 US-09-764-762-2	Sequence 2, Appl1
10	460	29.3	1504	3 US-09-280-116-1	Sequence 1, Appl1
11	253	16.1	601	4 US-09-949-016-63414	Sequence 63414, A
12	223.4	14.2	68719	4 US-09-949-016-12799	Sequence 12799, A
13	223.4	14.2	68720	4 US-09-949-016-14265	Sequence 14265, A
14	222.8	14.2	32278	4 US-09-949-016-14575	Sequence 14575, A
15	222.6	14.2	176373	3 US-09-128-155-17	Sequence 17, Appl
16	222.2	14.1	601	4 US-09-949-016-66914	Sequence 66914, A
17	222.2	14.1	177668	4 US-09-949-016-13713	Sequence 13713, A
18	221.8	14.1	80858	4 US-09-949-016-12659	Sequence 12659, A
19	221.8	14.1	80859	4 US-09-949-016-15715	Sequence 15715, A
20	220.8	14.1	69701	4 US-09-949-016-14187	Sequence 14187, A
21	220.8	14.1	73308	4 US-09-949-016-16336	Sequence 16336, A
22	219.8	14.0	300402	4 US-09-949-016-13632	Sequence 13632, A
23	219.4	14.0	38009	4 US-09-949-016-13617	Sequence 13617, A
24	219.4	14.0	70000	3 US-09-851-896-3	Sequence 3, Appl1
25	219.4	14.0	76399	4 US-09-949-016-16819	Sequence 16819, A
26	219.2	14.0	152331	3 US-09-128-155-16	Sequence 16, Appl
27	218.8	13.9	71574	4 US-09-949-016-15580	Sequence 15580, A

28	218.6	13.9	25464	4 US-09-326-480A-4	Sequence 4, Appl1
29	218.6	13.9	87039	4 US-09-949-016-15691	Sequence 15691, A
30	218.2	13.9	41755	4 US-09-949-016-15728	Sequence 15728, A
31	218	13.9	5695	4 US-09-949-016-13664	Sequence 13664, A
32	218	13.9	301828	4 US-09-949-016-13669	Sequence 13669, A
33	217.8	13.9	601	4 US-09-949-016-64323	Sequence 64323, A
34	217.8	13.9	36156	4 US-09-949-016-12128	Sequence 12128, A
35	217.8	13.9	36156	4 US-09-949-016-13261	Sequence 13261, A
36	217.8	13.9	46841	4 US-09-949-016-13466	Sequence 13466, A
37	217.6	13.9	246444	4 US-09-949-016-13113	Sequence 13113, A
38	217.6	13.9	321022	4 US-09-949-016-11852	Sequence 11852, A
39	217.6	13.9	321022	4 US-09-949-016-14166	Sequence 14166, A
40	217.2	13.8	101128	4 US-09-949-016-14293	Sequence 14293, A
41	216.8	13.8	10013	4 US-09-949-016-16474	Sequence 16474, A
42	216.6	13.8	44019	4 US-09-949-016-14902	Sequence 14902, A
43	216.6	13.8	461	3 US-09-404-879A-1	Sequence 1, Appl1
44	216.4	13.8	461	3 US-09-404-879A-3	Sequence 3, Appl1
45	216.4	13.8	461	4 US-09-338-933-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1	US-09-949-016-13583	
Sequence 13583	Application US/09949016	
Patent No. 6812338		
GENERAL INFORMATION:		
APPLICANT: VENTER, J. Craig et al.		
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED		
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF		
FILE REFERENCE: C10010307		
CURRENT APPLICATION NUMBER: US/09/949, 016		
CURRENT FILING DATE: 2000-04-14		
PRIOR APPLICATION NUMBER: 60/241, 755		
PRIOR FILING DATE: 2000-10-20		
PRIOR APPLICATION NUMBER: 60/237, 768		
PRIOR FILING DATE: 2000-10-03		
PRIOR APPLICATION NUMBER: 60/231, 498		
PRIOR FILING DATE: 2000-09-08		
NUMBER OF SEQ ID NOS: 207012		
SOFTWARE: FastSeq for Windows Version 4.0		
SEQ ID NO 13583		
LENGTH: 10818		
TYPE: DNA		
ORGANISM: Human		
US-09-949-016-13583		
Query Match	99.8% ; Score 1567.8 ; DB 4 ; Length 10818 ;	
Best Local Similarity	99.8% ; Pred. No. 0 ;	
Matches 1569 ; Conservative	0 ; Mismatches 2 ; Indels 0 ; Gaps 0 ;	
QY	1	AGGAGAGTATGGGGAATTGAGACAGAAACAAATTAGTCCAGCGAATGATTCTA 60
DB	7243	AGGAGAGTATGGGGAATTGAGACAGAAACAAATTAGTCCAGCGAATGATTCTA 7302
QY	61	TTGGGAGTATTTGTCCTCTAGAACACTGCGAATACCAAGAGACATTTTGGTTGTA 120
DB	7303	TTGGGAGTATTTGTCCTCTAGAACACTGCGAATACCAAGAGACATTTTGGTTGTA 7362
QY	121	CAACTATATGAGGGGCACTTACGCACTAATGATGATGCAAGTGGCTGTTCAAC 180
DB	7363	CAACTATATGAGGGGCACTTACGCACTAATGATGATGCAAGTGGCTGTTCAAC 7422
QY	181	ATGCTATGATGACACAGGAGGCTTACCAACAAACATTATTCAGATGCCAC 240
DB	7423	ATGCTATGATGACACAGGAGGCTTACCAACAAACATTATTCAGATGCCAC 7482
QY	241	AGTCCCAATGACAGAACCCCTATCAGAGGCTGGAACCGTATTTTGGCAGAGGAG 300
DB	7483	AGTCCCAATGACAGAACCCCTATCAGAGGCTGGAACCGTATTTTGGCAGAGGAG 7542
QY	301	GTTATAGATGGTGTGTGTGAGATGGGAGAGGTGTGTCTCAGTAAGAAATTA 360

```
Db 7543 GATAAAGATGGTGTGGAGAAATGGGAAAGGAAGTGTGTCCAGTAAGAGAAATTA 7602
Qy 361 GGCCTGACAGAGCTGGAGGGAGAGAGAGAGAGAGAGAGAGAGAGATACAGATGA 420
Db 7603 GGCCTGACAGAGCTGGAGGGAGAGAGAGAGAGAGAGAGAGAGAGATACAGATGA 7662
Qy 421 GGGAGACAGAGCTGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 7663 GGGAGACAGAGCTGGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7722
Qy 481 CAGACCCCTCCGAAATGATGTGTGACAAACAGGAATCTGGAAGAGAGAGAGAGAG 540
Db 7723 CAGACCCCTCCGAAATGATGTGTGACAAACAGGAATCTGGAAGAGAGAGAGAG 7782
Qy 541 AGTGAACAATGGGCTTAAGGTGAACCTTGAAGGACAGAGAGAGAGAGAGAGAGAG 600
Db 7783 AGTGAACAATGGGCTTAAGGTGAACCTTGAAGGACAGAGAGAGAGAGAGAGAG 7842
Qy 601 ATCCCAACACTTGAAGGCTGAGAGTGGGAGAAATCACTTGAAGGACAGAGAGAG 660
Db 7843 ATCCCAACACTTGAAGGCTGAGAGTGGGAGAAATCACTTGAAGGACAGAGAGAG 7902
Qy 661 CCTGGCAACATGTGTGAAAACCCGCTCTACAAAATAATACAAAATAATAGCCGG 720
Db 7903 CCTGGCAACATGTGTGAAAACCCGCTCTACAAAATAATACAAAATAATAGCC 7962
Qy 721 GGTATAGACACCTGATGACAGCTAATCTTGGAGGCTGAGAGAGAGAGAGAGAG 780
Db 7963 GGTATAGACACCTGATGACAGCTAATCTTGGAGGCTGAGAGAGAGAGAGAGAG 8022
Qy 781 CCGGAGATGAGAGCTGACAGTGAAGTCAAGGCTGAGGCTCAACTGAGGAGAGAG 840
Db 8023 CCGGAGATGAGAGCTGACAGTGAAGTCAAGGCTGAGGCTCAACTGAGGAGAGAG 8082
Qy 841 GAGTAAGACTCCATCTCAAAAAAAGCTGAGATTTGAGTGAATAATTAATTAAT 900
Db 8083 GAGTAAGACTCCATCTCAAAAAAAGCTGAGATTTGAGTGAATAATTAATTAAT 8142
Qy 901 CTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Db 8143 CTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8202
Qy 961 CTGTAATTTCAATCTCTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1020
Db 8203 CTGTAATTTCAATCTCTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 8262
Qy 1021 GGTCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db 8263 GGTCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8322
Qy 1081 TCTCTCTGCAAGGAGATCTGAGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 8323 TCTCTCTGCAAGGAGATCTGAGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 8382
Qy 1141 GTGTCTCTGAGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 8383 GTGTCTCTGAGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8442
Qy 1201 TGCAGATTCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 8443 TGCAGATTCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8502
Qy 1261 CTCAGACACGGGACATCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
Db 8503 CTCAGACACGGGACATCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8562
Qy 1321 CATTCCTTCCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1380
Db 8563 CATTCCTTCCAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8622
Qy 1381 GGAATCAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1440
```

```
Db 8623 GGAATCAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8682
Qy 1441 AACATTTTCCAAATGTCACAGGGGGGTTGGCTCAATCTCCCTGGGACATTTCA 1500
Db 8683 AACATTTTCCAAATGTCACAGGGGGGTTGGCTCAATCTCCCTGGGACATTTCA 8742
Qy 1501 TCTCAAGCTCAGGGGCCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAAT 1560
Db 8743 TCTCAAGCTCAGGGGCCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAAT 8802
Qy 1561 AAATGAGAG 1571
Db 8803 AAATGAGAG 8813

RESULT 2
US-09-949-016-63416/c
; Sequence 63416, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63416
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63416

Query Match 38.1%; Score 599; DB 4; Length 601;
Best Local Similarity 99.7%; Pred. No. 3.4e-151;
Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 482 AGACCCCTCCGAAATGATGTGTGACAAACAGGAATCTGGAAGAGAGAGAGAGAGAG 541
Db 601 AGACCCCTCCGAAATGATGTGTGACAAACAGGAATCTGGAAGAGAGAGAGAGAGAG 542
Qy 542 GTGACAAATGGGGTCTTAAAGTTGAATCTTGAAGGACAGGAGATGTGTGTGTGTGTGT 601
Db 541 GTGACAAATGGGGTCTTAAAGTTGAATCTTGAAGGACAGGAGATGTGTGTGTGTGTGT 482
Qy 602 TCCCAACACTTGTGAGGCTGAGGTGGGAGATCACTTGAAGGACAGAGAGTGTGAGAC 661
Db 481 TCCCAACACTTGTGAGGCTGAGGTGGGAGATCACTTGAAGGACAGAGAGTGTGAGAC 422
Qy 662 CTGGCAACATGAGTGAACCCCGTCTCAAAAAAATTAAGCCGGGGTGTG 721
Db 421 CTGGCAACATGAGTGAACCCCGTCTCAAAAAAATTAAGCCGGGGTGTG 362
Qy 722 GTGATGACACCTGATGACAGCTAATTTGAGAGGCTGAGGACAGAGAGATTTGAGAAC 781
Db 361 GTGATGACACCTGATGACAGCTAATTTGAGAGGCTGAGGACAGAGAGATTTGAGAAC 302
Qy 782 CGGAGATGAGAGGCTGACGTGAGCTGAGGCTCACTGGCTTCAACTGTGGCAACAG 841
Db 301 YGGAGATGAGAGGCTGACGTGAGCTGAGGCTCACTGGCTTCAACTGTGGCAACAG 242
Qy 842 AGTAAGACTCCATCTCAAAAAAAGCTGAGATTTGAGAGTGAATAATTAATTAATTC 901
Db 241 AGTAAGACTCCATCTCAAAAAAAGCTGAGATTTGAGAGTGAATAATTAATTAATTC 182
Qy 902 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 961
```



```
/ Sequence 1, Application US/09509908
/ Patent No. 6589770
/ GENERAL INFORMATION:
/ APPLICANT: The Procter & Gamble Company, N/A N/A
/ TITLE OF INVENTION: A Protease
/ NUMBER OF SEQUENCES: 2
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: T. David Reed
/ STREET: 5299 Spring Grove Avenue
/ CITY: Cincinnati
/ STATE: Ohio
/ COUNTRY: USA
/ ZIP: 45217-1087
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/509,908
/ FILING DATE: 28-Feb-2002
/ CLASSIFICATION: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Reed, T David
/ REGISTRATION NUMBER: 32,931
/ REFERENCE/DOCKET NUMBER: AA-264F
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 513-627-7025
/ TELEFAX: 513-627-6333
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1499 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 291..1172
/ FEATURE:
/ NAME/KEY: mat_peptide
/ LOCATION: 489..1172
/ FEATURE:
/ NAME/KEY: 5'UTR
/ LOCATION: 1..290
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: 1173..1499
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-509-908-1

Query Match      30.7%; Score 482; DB 4; Length 1499;
Best Local Similarity 100.0%; Pred.No.1.5e-119;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGATGATTCGGGGGCGCTGTGATGCTCAATGAGCTCCCTGAGAGGACTCGTGTCTGG 1149
DB 1014 CAGGGATGATTCGGGGGCGCTGTGATGCTCAATGAGCTCCCTGAGAGGACTCGTGTCTGG 1073
QY 1150 GGAAGATTACCTTTGTCGCCGCCCAACAGACCGGGGTGTCTACAGCAACTCTTCGAATTC 1209
DB 1074 GGAAGATTACCTTTGTCGCCGCCCAACAGACCGGGGTGTCTACAGCAACTCTTCGAATTC 1133
QY 1210 ACCAAGTGAATCCAGGAACCAATCCAGGCAACTCCTGAGTCAATCCAGAGACTCAGACAA 1269
DB 1134 ACCAAGTGAATCCAGGAACCAATCCAGGCAACTCCTGAGTCAATCCAGAGACTCAGACAA 1193
QY 1270 CCGGATCCCACTGCTGCTGAGGAGACAGCCCTGACACTCTTTCAGACCTCAATTCCTTC 1329
DB 1194 CCGGATCCCACTGCTGCTGAGGAGACAGCCCTGACACTCTTTCAGACCTCAATTCCTTC 1253
QY 1330 CCAGAGATGTGAAGATGTTCACTCTCCAGCCCTGACCCCAATGTTCTCTGAGACTCAGG 1389
```

```
DB 1254 CCAGAGATGTGAAGATGTTCACTCTCCAGCCCTGACCCCAATGCTCTCCGACTCAGG 1313
QY 1390 GTCGTCTTCCCACTGATGAGGCTGACCGGTCTCTCTAGTTGAACCTGAGAACATTTTC 1449
DB 1314 GTCGTCTTCCCACTGATGAGGCTGACCGGTCTCTCTAGTTGAACCTGAGAACATTTTC 1373
QY 1450 CAAAATGTCAGGGCGGGGGGTGGCTGCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 1509
DB 1374 CAAAATGTCAGGGCGGGGGGTGGCTGCTCAATCTCCCTGGGGCACTTTCATCTCAAGC 1433
QY 1510 TCAGGGCCCATCTCTTCTCTGAGCTGAGCCCAATTTAGTCCAGAAATTAATCTGAGA 1569
DB 1434 TCAGGGCCCATCTCTTCTCTGAGCTGAGCCCAATTTAGTCCAGAAATTAATCTGAGA 1493
QY 1570 AG 1571
DB 1494 AG 1495

RESULT 6
US-09-949-016-63415/c
/ Sequence 63415, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: C1001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 63415
/ LENGTH: 601
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-63415

Query Match      30.6%; Score 480.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred.No.2.3e-119;
Matches 480; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGATATGGGGAATTGAAGACAGAAACAAATTAATGTCAGGAAATGATTTCTA 60
DB 481 AGGAGGATATGGGGAATTGAAGACAGAAACAAATTAATGTCAGGAAATGATTTCTA 422
QY 61 TTGGAGATGATTTTCCCTTAGAAGACACTGGCAATACCAGAGACATTTTGTGTGCA 120
DB 421 TTGGAGATGATTTTCCCTTAGAAGACACTGGCAATACCAGAGACATTTTGTGTGCA 362
QY 121 CAACTATATGAGGGGCAATTAATGCACTAATGATGATGATGATGATGATGATGATGATGAT 180
DB 361 CAACTATATGAGGGGCAATTAATGCACTAATGATGATGATGATGATGATGATGATGATGAT 302
QY 181 ATGCTATGATGACACGCGACAGGCTTCCACAAACCAATTAATGATGATGATGATGATGATGAT 240
DB 301 ATGCTATGATGACACGCGACAGGCTTCCACAAACCAATTAATGATGATGATGATGATGATGAT 242
QY 241 AGTGCCCATGATGAGAAACCTTCATCCAGGGGCTGAGAAACGTAATTTTTCAGAGGGAG 300
DB 241 AGTGCCCATGATGAGAAACCTTCATCCAGGGGCTGAGAAACGTAATTTTTCAGAGGGAG 182
QY 301 GATATAGATGGGTTGTGAGAAATGGGAGAAAGTGTGTGTCAGTAAAGAAATTA 360
DB 181 GATATAGATGGGTTGTGAGAAATGGGAGAAAGTGTGTGTCAGTAAAGAAATTA 122
QY 361 GGCTTGACAGGCTTGAAGGGGAGAGTGAAGAGAAAGGAGCGGAGAGATACAGATGA 420
```


Db 121 GGGCTGCACAGGCTGGAGGGGAGAGTGAAGAAAGGAGCGGAGAGATACAGATGA 62
Qy 421 GGGAGACAGGCTGGAAACAGAAAGTGAAGACGAATTCAGAGATGGAGAGAAAGGCTCA 480
Db 61 GGGAGACAGGCTGGAAACAGAAAGTGAAGACGAATTCAGAGATGGAGAGAAAGGCTCA 2
Qy 481 C 481
Db 1 C 1

RESULT 7
US-08-824-874-2
Sequence 2, Application US/08824874
Patent No. 5962300
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE: Filed Herewith
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-08-824-874-2

Query Match 30.3%; Score 476; DB 2; Length 1476;
Best Local Similarity 100.0%; Pred. No. 66-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCGTCAGGGAGCTGTCTGG 1149
Db 1001 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCGTCAGGGAGCTGTCTGG 1060
Qy 1150 GGAATTAACCTTGTGCGCGGCGCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db 1061 GGAGATTACCTTGTGCGCGGCGCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1120
Qy 1210 ACCAAGTGAATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 1269
Db 1121 ACCAAGTGAATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 1180

Qy 1270 CCGGCAATCCCACTGTGTGAGGGAGACGCCCTGAACAATCCTTTCAGACCTCAATCCTTC 1329
Db 1181 CCGGCAATCCCACTGTGTGAGGGAGACGCCCTGAACAATCCTTTCAGACCTCAATCCTTC 1240
Qy 1330 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGAGCCCAATGTTCTCTGAGACTCAGG 1389
Db 1241 CCAGAGATGTTGAAATGTTCAATCTCTCCAGCCCTGAGCCCAATGTTCTCTGAGACTCAGG 1300
Qy 1390 GTCTGCTTCCCAATGTTGAGCTGACCGGTGTCTCTGATTGAACCTTGGAACAATTTTC 1449
Db 1301 GTCTGCTTCCCAATGTTGAGCTGACCGGTGTCTCTGATTGAACCTTGGAACAATTTTC 1360
Qy 1450 CAAAACGTCCAGGGGGGGGCTGCGTCAATCTCCGCGGACATTCATCTCAAGC 1509
Db 1361 CAAAACGTCCAGGGGGGGGCTGCGTCAATCTCCGCGGACATTCATCTCAAGC 1420
Qy 1510 TCAGGGCCCATCCCTTCTGTGACGCTTGACCCCAATTTAGTCCAGAAATAAAT 1565
Db 1421 TCAGGGCCCATCCCTTCTGTGACGCTTGACCCCAATTTAGTCCAGAAATAAAT 1476

RESULT 8
US-09-210-084-2
Sequence 2, Application US/09210084
Patent No. 6197511
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Lal, Preeti
TITLE OF INVENTION: NOVEL KALLIKREIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-09-210-084-2

Query Match 30.3%; Score 476; DB 3; Length 1476;
Best Local Similarity 100.0%; Pred. No. 66-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1090 CAGGGTATCTGGGGGCGCTGTGTGCAATGGCTCCGTCAGGGAGCTGTCTGG 1149

DB 1001 CAGGGGATGATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTCTCG 1060
QY 1150 GGAGATTACCTCTGTGCCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 1061 GGAGATTACCTCTGTGCCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1120
QY 1210 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGGACTCAGACACA 1269
DB 1121 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGGACTCAGACACA 1180
QY 1270 CCGGCAATCCCACTGCTGCAAGGAAACAGCCCTGACACTCTCTTCAAGACCTCAATCTCTTC 1329
DB 1181 CCGGCAATCCCACTGCTGCAAGGAAACAGCCCTGACACTCTCTTCAAGACCTCAATCTCTTC 1240
QY 1330 CCAGAGATGTTGGAATGTTATCTCTGCAAGCCCGCCAGACCCCAATGTTCTCTGAGACTCAG 1389
DB 1241 CCAGAGATGTTGGAATGTTATCTCTGCAAGCCCGCCAGACCCCAATGTTCTCTGAGACTCAG 1300
QY 1390 GTCTGCTTCCCACTGCTGCAAGGAAACAGCCCTGATGTATCCAGGAACTCTGCAAGTTC 1449
DB 1301 GTCTGCTTCCCACTGCTGCAAGGAAACAGCCCTGATGTATCCAGGAACTCTGCAAGTTC 1360
QY 1450 CAAAATGTCAGGCGGGGGTGTGCTGCAATCTCTGAGGCACTTTCATCTCAAGC 1509
DB 1361 CAAAATGTCAGGCGGGGGTGTGCTGCAATCTCTGAGGCACTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1565
DB 1421 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1476

RESULT 9

US-09-764-762-2

; Sequence 2, Application US/09764762

; Patent No. 6472195

; GENERAL INFORMATION:

; APPLICANT: Hallman, Jennifer L.

; Inventor: Hall, Preeti

; TITLE OF INVENTION: NOVEL KALLIKREIN

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/764,762

; FILING DATE: 16-Jan-2001

; CLASSIFICATION: <Unknown>

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 09/210,084

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0252 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1476 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 820694
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Query Match 30.3%; Score 476; DB 4; Length 1476;
Best Local Similarity 100.0%; Pred. No. 6e-118;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTCTCG 1149
DB 1001 CAGGGTGAATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTCTCG 1060
QY 1150 GGAGATTACCTCTGTGCCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 1061 GGAGATTACCTCTGTGCCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1120
QY 1210 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGGACTCAGACACA 1269
DB 1121 ACCAAGTGTATCCAGGAAACCATTCAGGCAACTCTGATGTATCCAGGACTCAGACACA 1180
QY 1270 CCGGCAATCCCACTGCTGCAAGGAAACAGCCCTGACACTCTCTTCAAGACCTCAATCTCTTC 1329
DB 1181 CCGGCAATCCCACTGCTGCAAGGAAACAGCCCTGACACTCTCTTCAAGACCTCAATCTCTTC 1240
QY 1330 CCAGAGATGTTGGAATGTTATCTCTGCAAGCCCGCCAGACCCCAATGTTCTCTGAGACTCAG 1389
DB 1241 CCAGAGATGTTGGAATGTTATCTCTGCAAGCCCGCCAGACCCCAATGTTCTCTGAGACTCAG 1300
QY 1390 GTCTGCTTCCCACTGCTGCAAGGAAACAGCCCTGATGTATCCAGGAACTCTGCAAGTTC 1449
DB 1301 GTCTGCTTCCCACTGCTGCAAGGAAACAGCCCTGATGTATCCAGGAACTCTGCAAGTTC 1360
QY 1450 CAAAATGTCAGGCGGGGGTGTGCTGCAATCTCTGAGGCACTTTCATCTCAAGC 1509
DB 1361 CAAAATGTCAGGCGGGGGTGTGCTGCAATCTCTGAGGCACTTTCATCTCAAGC 1420
QY 1510 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1565
DB 1421 TCAGGGCCCATCCCTTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAAC 1476

RESULT 10

US-09-280-116-1

; Sequence 1, Application US/09280116A

; Patent No. 6331427

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs

; FILE REFERENCE: 5800-24, 035800/176965

; CURRENT APPLICATION NUMBER: US/09/280,116A

; CURRENT FILING DATE: 1999-03-26

; NUMBER OF SEQ ID NOS: 268

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1504

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: trypsin-like serine proteases

; US-09-280-116-1

Query Match 29.3%; Score 460; DB 3; Length 1504;
Best Local Similarity 99.6%; Pred. No. 1.2e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1090 CAGGGTGAATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTCTCG 1148
DB 1015 CAGGGTGAATTCGGGGGCGCTGTGCTGCAATGGCTCCCTGAGGAACTCGTCTCG 1074
QY 1149 GGAGATTACCTTGTGCCCCGAGCAAGACCGGGGTGTCTACAGAACTCTGCAAGTT 1208

Db	1075	GGGAGATTACCTTGTGCCCCGGGCCAACAAGACCGGGTGTCTAACAGAACTCTGCAAGTT	1134
OY	1209	CAGCAGATGGAATCCAGAGAAACCATTCAGAGCCAACTCTGAAGTCATCCAGGACTCAGCAC	1268
Db	1135	CAGCAAGTGAATCCAGAGAAACATTCAGAGCCAACTCTGAAGTCATCCAGGACTCAGCAC	1194
OY	1269	ACGGGCAATCCCACTCTGTCGAGGGACAAGCCTTGACACTCTTTTCAGACCTTCATTCTTT	1328
Db	1195	ACGGGCAATCCCACTCTGTCGAGGGACAAGCCTTGACACTCTTTTCAGACCTTCATTCTTT	1254
OY	1339	CCGAGAGATGTTGAGAAATGTTCAATCTTCTCCAGCCCTTACCCCAATGTTCTCTGAGACTCAG	1388
Db	1255	CCGAGAGATGTTGAGAAATGTTCAATCTTCTCCAGCCCTTACCCCAATGTTCTCTGAGACTCAG	1314
OY	1389	GGTCGCTTCCCCCAATTGGGCTGACCGTGTCTCTCTAAGTTGAACCTTGGGAAACAATTT	1448
Db	1315	GGTCGCTTCCCCCAATTGGGCTGACCGTGTCTCTCTAAGTTGAACCTTGGGAAACAATTT	1374
OY	1449	CCAAACATGTCAGAGGCGGGGGGTGGCGTCAATCCCTGGGGACAATTATCCTCAAG	1508
Db	1375	CCAAACATGTCAGAGGCGGGGGGTGGCGTCAATCCCTGGGGACAATTATCCTCAAG	1434
OY	1509	CTGAGGGGCCATCCCTTCTCTGACGCTCGAACCCAAATTTAGT--CCCGAATAATACTGA	1567
Db	1435	CTGAGGGGCCATCCCTTCTCTGACGCTCGAACCCAAATTTAGTCCCGAATAATACTGA	1494
OY	1568	GAAG 1571	
Db	1495	GAAG 1498	

```

RESULT 11
US-09-949-016-63414/C
; Sequence 63414, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241, 755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237, 768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231, 498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63414
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63414

```

Query Match %	16.1%	Score 253	DB 4	Length 601
Best Local Similarity	100.0%	Pred.NO.	3.5e-58	
Matches 253, Conservative	0	Mismatches	0	Indels 0
				Gaps 0

QY	1	AGGGAGGTATGGGGAATTGAAAGCAGAAACAAATTAGTCCAAAGGAAATGATTTCTA	60
Db	253	AGGGAGGTATGGGGAATTGAAAGCAGAAACAAATTAGTCCAAAGGAAATGATTTCTA	194
QY	61	TTGGGAGTGATTTCTGCCCCTAGAAGCACATGGCAATACAGAGACATTTTGGTGTCA	120
Db	193	TTGGGAGTGATTTCTGCCCCTAGAAGCACATGGCAATACAGAGACATTTTGGTGTCA	134
QY	121	CAACTATATGAGAGGGCAATTACTGGCAACTAATGATAGATGCGAAGTGTGCTTCAAC	180
Db	133	CAACTATATGAGAGGGCAATTACTGGCAACTAATGATAGATGCGAAGTGTGCTTCAAC	74
QY	181	ATGCTATATGACACAGCAGGCGCTCCACAACAACCATTTATCCAGCTTCAGATGCCAC	240

Db 73 ATGCTTATGATGCGACACGGCAGGCGCTCCAAACAAACCATTCACAGTCCAC 14
|||
QY 241 AGTGCCCGATCG 253
|||
Db 13 AGTGCCCGATCG 1
|||

```

RESULT 12
US-09-949-016-12799
: Sequence 12799, Application US/03949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: WITH INFECTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C0001307
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 12799
: LENGTH: 68719
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-12799

```

Query Match	14.2%	Score 223.4	DB 4	Length 68719
Best Local Similarity	86.0%	Pred. No. 2.8e-49		
Matches 259	Conservative	0	Mismatches 4	Indels 1
				Gaps 1

QY	570	TTGAGGCGACGCAATGGTGGCTCAGGCTGTGAATATCCCAACATTT--GGAAGCTGAGGTGG	628
Db	57314	TGAGGCTGGGCAACGGTGGCTCAAGCTGTAAATCCAGCACTTTGGGAGGCCAAGCAGG	57373
QY	629	CGAATCACTTTGAGGCCAGAGATTGAGACCAAGCTTGCCCAATGGTGAACCCGCTTC	688
Db	57374	TGGATCACTCGAGGCTCAGGAGTTTCAAGACCAAGCTGGCCCAATGGCGAAACCCGCTTC	57433
QY	689	TACAAAAAAAATACAAAAAATTGACCGGGGNGTGGTGAATGACACCTGTGTGCACAGCTTC	748
Db	57434	TACTTAAAAAATACAGAAAAATTTAGCCAGGCAATGTGTGGGCACTGTAAATCCAGTGC	57493
QY	749	TTGGAGGCTGAGGCGAGAGAAATTGCTTGAACCCGGAGATGAGGCTGCAGTGAAGCTGA	808
Db	57494	TTGGGAGGCTGAGGCGAGAGAAATTGCTTGAACCTGGAGACAGAGGTTGTCAATGAGCTGA	57553
QY	809	GGTCAGGCGCACTCGCTCCAACTTGGGCAACAGATTPAAGTCTCAATCTCAAAAAAAAA	868
Db	57554	GATGTGTCCTCATTTACTTCCACTCGGCGCAACAGAGGGAATCTCCCTTCAGAAAAAAAA	57613
QY	869	A 869	
Db	57614	A 57614	

```

RESULT 13
US-09-949-016-14296
; Sequence 14296, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949.016
; CURRENT FILING DATE: 2000-04-14

```

PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14296
LENGTH: 68720
TYPE: DNA
ORGANISM: Human
US-09-949-016-14296

Query Match 14.2%; Score 223.4; DB 4; Length 68720;
Best Local Similarity 86.0%; Pred. No. 2.8e-49;
Matches 259; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 570 TGGAGGCCAGGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GGAGGCTGAGGTGG 628
DB 57314 TGGAGGCTGGGCACGGTGGCTCAAGCTGTATCCAGACACTTTGGGAGGCCAAGGCAAG 57373
QY 629 CGATCACTTGAGGCCAGAGTTTCAGACCAAGCTTGCCCAACATGTGAAACCCGCTTC 688
DB 57374 TGGATCACTGAGGTCAAGAGTTCAAGACCAAGCTTGCCCAACATGTGCAAAACCCGCTTC 57433
QY 689 TACAAAAAATACAAAAATTAGCCGGGTGTGTATGAGACCTGTGTCTACAGACTAC 748
DB 57434 TACTAAAAAATACAAAAATTAGCCAGCATGTGTGTGGACCTGTATTCCTCAAGTGA 57493
QY 749 TTGGAGGCTGAGCAGAGAAATGTGTAACCCGGAGATGAGGCTGACAGTGA 808
DB 57494 TTGGAGGCTGAGCAGAGAAATGTGTAACCTGTGGAACCTGGAGACAGAGTTGACAGTGA 57553
QY 809 GGTGAGGCCACTCGCTCCCACTGGGCAACAGATGAAGTCAATCTCAAAAAA 868
DB 57554 GATCGGCCACTGTACTCCACCTGGGCAACAGAGAGGACTCTCTCAGAAAAA 57613
QY 869 A 869
DB 57614 A 57614

RESULT 14
US-09-949-016-14575
Sequence 14575, Application US/09949016
Patent No. 6812338
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14575
LENGTH: 32278
TYPE: DNA
ORGANISM: Human
US-09-949-016-14575

Query Match 14.2%; Score 222.8; DB 4; Length 32278;
Best Local Similarity 85.8%; Pred. No. 2.9e-49;
Matches 259; Conservative 0; Mismatches 42; Indels 1; Gaps 1;
QY 569 TTGGAGGCCAGGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GGAGGCTGAGGTG 627

DB 13360 TTGTTGGCCAGGCGTAGTGGCTCAGCGCTGTATCCAGACACTTTGGGAGACCAAGTGG 13419
QY 628 GGAATCACTTAGGCCAGAGATTGGAGACCAAGCTTGCCCAACATGTGAAACCCGCT 687
DB 13420 GGAGATCACTTAAGGTCAAGAGATTGAGACCAAGCTTGCCCAACATGTGAAACCCGCT 13479
QY 688 CTACAAAAAATACAAAAAATTAGCCGGGTGTGTGATGACACTGTAGTCAAGACTA 747
DB 13480 CTACAAAAAATACAAAAAATTAGCCAGGTGTGTGTGAGAGTGTGCTGTATCCAGCTA 13539
QY 748 CTTGGAGGCTGAGGAGAGAAATGCTTGAACCCGGAGATGAGAGCTGCACTGAGCTG 807
DB 13540 CTTGGAGGCTGAGGAGAGAAATGCTTGAACCCGGAGATGAGAGCTGCACTGAGCTG 13599
QY 808 AGTCAGGCGCACTGCGCTCCCACTGGGCAACAGATGAGACTCCATCCAAAAA 867
DB 13600 AGATGCGCCACTGCACTCCAGCTGGGCAACAGAGGAGACTCCGCTCAAAACAAAC 13659
QY 868 AA 869
DB 13660 AA 13661

RESULT 15
US-09-128-155-17
Sequence 17, Application US/09128155
Patent No. 6117654
GENERAL INFORMATION:
APPLICANT: Pan, Yang
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
AND USES THEREOF
FILE REFERENCE: 09404/052001
CURRENT APPLICATION NUMBER: US/09/128,155
CURRENT FILING DATE: 1998-08-03
EARLIER APPLICATION NUMBER: US 60/091,650
EARLIER FILING DATE: 1998-07-02
EARLIER APPLICATION NUMBER: US 60/054,646
EARLIER FILING DATE: 1997-08-04
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 176373
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(176373)
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match 14.2%; Score 222.6; DB 3; Length 176373;
Best Local Similarity 84.0%; Pred. No. 7.2e-49;
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;
QY 558 AAGGTTGAAGTTGAGGCCAGGCATGTGTGCTCAGCGCTGTATCCCAACACTTT-GGA 616
DB 127012 AAGAAAAAATTAATTTGGCAGGCAAGGCGGTCTATGCTTATATCCAGCACTTTGGGA 127071
QY 617 GGTGAGGTGGGCAATCACTTGAGGCCAGAGTTTCAGACCAAGCTTGCCCAACATGTG 676
DB 127072 GGTGAGGAGGTGATCACTGAGGTCAAGAGTTTCAGACCAAGCTTGCCCAACATGTG 127131
QY 677 AAACCCGCTCTACAAAAAATACAAAAAATTAGCCGGGTGTGTATGAGCACTGT 736
DB 127132 AAACCCGCTCTCTACAAAAAATACAAAAAATTAGCCAGGTGTGTGTGAGCACTGT 127191
QY 737 AGTCAGGCTACTTGGAGGCTGAGGAGAGAAATGCTTGAACCCGGAGATGAGGCT 796
DB 127192 AATCCAGCTACAGGAGGCTGAGGAGAGAAATGCTTGAACCCGGAGAGAAAT 127251
QY 797 GAGTGAAGTGAAGTCAAGGCCACTGCTCCAACTGGGCAACAGATGAAGCTCACT 856

Db 127252 GCAGTGAGCCAGTTCAAGCCACTGCACCCGCGCTGGGCAACGAGCAAGACTTTGTCT 127311
Qy 857 CAAAAAAAAAAAA 869
Db 127312 CAAAAAAAAAAAA 127324

Search completed: February 27, 2005, 05:14:39
Job time : 250.117 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:12 ; Search time 845.613 Seconds
(without alignments)
11008.461 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 agsgaggtacgaggaattga.....ccagaataactgagag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/1/pubpna/US10F_PUBCOMB.seq:*
- 20: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
- 22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	541	34.4	567	16	US-10-029-386-7604 Sequence 7604, Ap
2	482	30.7	586	9	US-09-954-531-849 Sequence 849, App
3	482	30.7	586	9	US-09-954-531-1266 Sequence 1266, Ap
4	482	30.7	735	9	US-09-745-288-94 Sequence 94, Appl
5	482	30.7	735	17	US-10-453-919-94 Sequence 94, Appl
6	482	30.7	1260	17	US-10-172-118-1411 Sequence 1411, Ap
7	482	30.7	1260	17	US-10-295-027-515 Sequence 515, Appl
8	482	30.7	1260	17	US-10-173-999-53 Sequence 53, Appl
9	482	30.7	1260	17	US-10-342-887-1411 Sequence 1411, Ap
10	482	30.7	1381	9	US-09-739-907-52 Sequence 52, Appl
11	482	30.7	1381	11	US-09-938-671-52 Sequence 52, Appl

12	482	30.7	1439	9	US-09-739-907-53 Sequence 53, Appl
13	482	30.7	1439	11	US-09-938-671-53 Sequence 53, Appl
14	482	30.7	1481	18	US-10-723-860-5032 Sequence 5032, Ap
15	482	30.7	1516	9	US-09-739-907-37 Sequence 37, Appl
16	482	30.7	1516	11	US-09-938-671-37 Sequence 37, Appl
17	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
18	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
19	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
20	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
21	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
22	482	30.7	1570	9	US-09-989-722-308 Sequence 308, App
23	482	30.7	1570	9	US-09-991-073-308 Sequence 308, App
24	482	30.7	1570	9	US-09-990-442-308 Sequence 308, App
25	482	30.7	1570	9	US-09-991-163-308 Sequence 308, App
26	482	30.7	1570	9	US-09-993-604-308 Sequence 308, App
27	482	30.7	1570	9	US-09-990-456-308 Sequence 308, App
28	482	30.7	1570	9	US-09-989-721-308 Sequence 308, App
29	482	30.7	1570	9	US-09-992-598-308 Sequence 308, App
30	482	30.7	1570	9	US-09-989-293A-308 Sequence 308, App
31	482	30.7	1570	9	US-09-989-735-308 Sequence 308, App
32	482	30.7	1570	9	US-09-990-444-308 Sequence 308, App
33	482	30.7	1570	9	US-09-991-181-308 Sequence 308, App
34	482	30.7	1570	9	US-09-989-730-308 Sequence 308, App
35	482	30.7	1570	9	US-09-990-436-308 Sequence 308, App
36	482	30.7	1570	9	US-09-993-687-308 Sequence 308, App
37	482	30.7	1570	10	US-09-989-734-308 Sequence 308, App
38	482	30.7	1570	10	US-09-997-653-308 Sequence 308, App
39	482	30.7	1570	10	US-09-989-724-308 Sequence 308, App
40	482	30.7	1570	10	US-09-989-728-308 Sequence 308, App
41	482	30.7	1570	10	US-09-990-441-308 Sequence 308, App
42	482	30.7	1570	10	US-09-993-667-308 Sequence 308, App
43	482	30.7	1570	10	US-09-997-422-308 Sequence 308, App
44	482	30.7	1570	10	US-09-997-666-308 Sequence 308, App
45	482	30.7	1570	10	US-09-990-438-308 Sequence 308, App

ALIGNMENTS

RESULT 1
US-10-029-386-7604
; Sequence 7604, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7604
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
; OTHER INFORMATION: EST HUMAN HIT: W73140.1, EVALUATE 0.00e+00
; OTHER INFORMATION: NT HIT: AP135028.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y337, EVALUATE 6.00e-27
US-10-029-386-7604

Query Match 34.4%; Score 541; DB 16; Length 567;
Best Local Similarity 100.0%; Pred. No. 7.2e-150;

Matches	541; Conservative	0; Mismatches	0; Indels	0; Gaps	0;
QY	1031	GTCTCTCTTCTGCCCACATTTTGCCACATCTCTGCTCTCTCTCATGCCCCCTTTTCTCTCTG	1090		
Db	1	GTCTCTCTTCTGCCCACATTTTGCCACATCTCTCTCTCTCATGCCCCCTTTTCTCTCTG	60		
QY	1091	AGGGTGATTTCTGGGGGGGCTGTGGTCTGCAATGAGCTCCCTGACAGGAATCGTGTCGAGG	1150		
Db	61	AGGGTGATTTCTGGGGGGGCTGTGGTCTGCAATGAGCTCCCTGACAGGAATCGTGTCGAGG	120		
QY	1151	GAGATTAACCTTGTGCCCCGCGCCCAACAGACCAGGCTGTCTACAGAACTCTGACAGTTCA	1210		
Db	121	GAGATTAACCTTGTGCCCCGCGCCCAACAGACCAGGCTGTCTACAGAACTCTGACAGTTCA	180		
QY	1211	CCAAATGATTCACAGAAACCATTCAGGCGCAACTCTGATGATCTCCAGAACTCAGCAC	1270		
Db	181	CCAAATGATTCACAGAAACCATTCAGGCGCAACTCTGATGATCTCCAGAACTCAGCAC	240		
QY	1271	CGGATATCCCAACCTGCTGACAGGAACGCCCTGACACTCTTTCAGAACCTCATTCCTCC	1330		
Db	241	CGGATATCCCAACCTGCTGACAGGAACGCCCTGACACTCTTTCAGAACCTCATTCCTCC	300		
QY	1331	CAGAGATTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGCTCTCTGAGACTCAGG	1390		
Db	301	CAGAGATTTGAGATGTTTCATCTCTCAGGCCCTGACCCCATGCTCTCTGAGACTCAGG	360		
QY	1391	TCTGCTTCTCCCAATTTGGGCTGACCGGTCTCTCTGATGTAACCTTGGGAACAATTTCC	1450		
Db	361	TCTGCTTCTCCCAATTTGGGCTGACCGGTCTCTCTGATGTAACCTTGGGAACAATTTCC	420		
QY	1451	AAACGTGTCACAGGCGGGGGGTGTGGCTCATCTCCCGGGGCACTTTCATCTCAAGCT	1510		
Db	421	AAACGTGTCACAGGCGGGGGGTGTGGCTCATCTCCCGGGGCACTTTCATCTCAAGCT	480		
QY	1511	CAGGCGCCATCTCTCTGCACTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGAA	1570		
Db	481	CAGGCGCCATCTCTCTGCACTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGAA	540		
QY	1571	G 1571			
Db	541	G 541			
RESULT 2					
US-09-954-531-849/c					
; Sequence 849, Application US/09954531					
; Patent No. US20020165180A1					
; GENERAL INFORMATION:					
APPLICANT: Weaver, Zoe					
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc					
FILE REFERENCE: 689290-77					
CURRENT APPLICATION NUMBER: US/09/954,531					
CURRENT FILING DATE: 2002-05-02					
PRIOR APPLICATION NUMBER: US/60/233,133					
PRIOR FILING DATE: 2000-09-18					
PRIOR APPLICATION NUMBER: US/60/234,009					
PRIOR FILING DATE: 2000-09-20					
PRIOR APPLICATION NUMBER: US/60/234,034					
PRIOR FILING DATE: 2000-09-20					
PRIOR APPLICATION NUMBER: US/60/234,509					
PRIOR FILING DATE: 2000-09-22					
PRIOR APPLICATION NUMBER: US/60/234,567					
PRIOR FILING DATE: 2000-09-22					
NUMBER OF SEQ ID NOS: 1392					
SOFTWARE: PatentIn version 3.0					
SEQ ID NO 849					
LENGTH: 586					
TYPE: DNA					
ORGANISM: Homo sapiens					
FEATURE:					
NAME/KEY: misc_feature					
LOCATION: (1)...(586)					

```

: OTHER INFORMATION: n=a,t,g or c
: US-09-954-531-849
:
: Query Match 30.7%; Score 482; DB 9; Length 586;
: Best Local Similarity 100.0%; Pred. No. 2,5e-132;
: Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
:
: QY 1090 CAGGATATTCCTGGGGGGCCCTGTGTGTGAAATGGCTCCCTGAGAGGACTCGTCTCTGG 1149
: Db 501 CAGGATATTCCTGGGGGGCCCTGTGTGTGAAATGGCTCCCTGAGAGGACTCGTCTCTGG 442
: QY 1150 GGAATTAACCTTGTGTGCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTCGAATTC 1209
: Db 441 GGAATTAACCTTGTGTGCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTCGAATTC 382
: QY 1210 ACCAATGATTCACAGAAACCATTCAGGCCCACTCTCTAGTCAATCCAGACTCAGACA 1269
: Db 381 ACCAATGATTCACAGAAACCATTCAGGCCCACTCTCTAGTCAATCCAGACTCAGACA 322
: QY 1270 CCGGATCCCCACTCTGTGTGAGGGACAGCCCTTGACACTCTTTTCAGACCTCTCAATCTTC 1329
: Db 321 CCGGATCCCCACTCTGTGTGAGGGACAGCCCTTGACACTCTTTTCAGACCTCTCAATCTTC 262
: QY 1330 CCAGAGATGTTGAAATGTTCAATCTCTCAGCCCTGACCCCAATGTTCTCTGACTCAGG 1389
: Db 261 CCAGAGATGTTGAAATGTTCAATCTCTCAGCCCTGACCCCAATGTTCTCTGACTCAGG 202
: QY 1390 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGTGGAACAATTTTC 1449
: Db 201 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTAGTTGAACCTGTGGAACAATTTTC 142
: QY 1450 CAAATCTGTCCAGGGCGGGGGGTTCGCTCTCAATCTCCCTGTGGGACATTTATCTCAAGC 1509
: Db 141 CAAATCTGTCCAGGGCGGGGGGTTCGCTCTCAATCTCCCTGTGGGACATTTATCTCAAGC 82
: QY 1510 TCAGGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 1569
: Db 81 TCAGGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 22
:
: QY 1570 AG 1571
: ||
: Db 21 AG 20
:
: RESULT 3
: US-09-954-531-1266/c
: Sequence 1266, Application US/09954531
: Patent No. US20020165180A1
: GENERAL INFORMATION:
: APPLICANT: Weaver, Zoe
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
: FILE REFERENCE: 689290-77
: CURRENT APPLICATION NUMBER: US/09/954,531
: PRIOR FILING DATE: 2002-05-02
: PRIOR APPLICATION NUMBER: US/60/233,133
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,009
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,034
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,509
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: US/60/234,567
: PRIOR FILING DATE: 2000-09-22
: NUMBER OF SEQ ID NOS: 1392
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 1266
: LENGTH: 586
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature

```


LOCATION: (1)...(586)
OTHER INFORMATION: n=a,t,g or c
US-09-954-531-1266

Query Match 30.7%; Score 482; DB 9; Length 586;
Best Local Similarity 100.0%; Pred. No. 2.5e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 1149
DB 501 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 442
QY 1150 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 441 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 1269
DB 381 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 322
QY 1270 CCGGATCCCACTCTGCTGCAAGGACAGCCCTGACACTCTTTTCAAGCCTCATCTTC 1329
DB 321 CCGGATCCCACTCTGCTGCAAGGACAGCCCTGACACTCTTTTCAAGCCTCATCTTC 262
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCAGCGCCCTGACCCCATGTCCTGAGCTCAG 1389
DB 261 CCAGAGATGTTGAGATGTTTCATCTCTCAGCGCCCTGACCCCATGTCCTGAGCTCAG 202
QY 1390 GTCTGCTTCCCACTGCTGCAAGGACAGCCCTGCTCTCTAGTTGAACTCTGGAACAATTC 1449
DB 201 GTCTGCTTCCCACTGCTGCAAGGACAGCCCTGCTCTCTAGTTGAACTCTGGAACAATTC 142
QY 1450 CAAAACGTCTCAGGGGCGGGGTTGCGTCTCATATCTCCCTGGGACATTTTCACTCAAG 1509
DB 141 CAAAACGTCTCAGGGGCGGGGTTGCGTCTCATATCTCCCTGGGACATTTTCACTCAAG 82
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 1569
DB 81 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 4
US-09-745-288-94/c
Sequence 94, Application US/09745288
Patent No. US20010018058A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
FILE REFERENCE: 210121.446D1
CURRENT APPLICATION NUMBER: US/09/745.288
CURRENT FILING DATE: 2000-12-19
NUMBER OF SEQ ID NOS: 101
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 94
LENGTH: 735
TYPE: DNA
ORGANISM: Homo sapien
US-09-745-288-94

Query Match 30.7%; Score 482; DB 9; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 1149
DB 500 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 441

QY 1150 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 440 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 381
QY 1210 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 1269
DB 380 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 321
QY 1270 CCGGATCCCACTCTGCTGCAAGGACAGCCCTGACACTCTTTTCAAGCCTCATCTTC 1329
DB 320 CCGGATCCCACTCTGCTGCAAGGACAGCCCTGACACTCTTTTCAAGCCTCATCTTC 261
QY 1330 CCAGAGATGTTGAGATGTTTCATCTCTCAGCGCCCTGACCCCATGTCCTCTGAGCTCAG 1389
DB 260 CCAGAGATGTTGAGATGTTTCATCTCTCAGCGCCCTGACCCCATGTCCTCTGAGCTCAG 201
QY 1390 GTCTGCTTCCCACTGCTGCAAGGACAGCCCTGCTCTCTAGTTGAACTCTGGAACAATTC 1449
DB 200 GTCTGCTTCCCACTGCTGCAAGGACAGCCCTGCTCTCTAGTTGAACTCTGGAACAATTC 141
QY 1450 CAAAACGTCTCAGGGGCGGGGTTGCGTCTCATATCTCCCTGGGACATTTTCACTCAAG 1509
DB 140 CAAAACGTCTCAGGGGCGGGGTTGCGTCTCATATCTCCCTGGGACATTTTCACTCAAG 81
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 1569
DB 80 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTTGTCCAGAAATTAACCTGAGA 21
QY 1570 AG 1571
DB 20 AG 19

RESULT 5
US-10-453-919-94/c
Sequence 94, Application US/10453919
Publication No. US20040033230A1
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Retter, Marc W.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.446C7
CURRENT APPLICATION NUMBER: US/10/453.919
CURRENT FILING DATE: 2003-06-03
NUMBER OF SEQ ID NOS: 121
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 94
LENGTH: 735
TYPE: DNA
ORGANISM: Homo sapien
US-10-453-919-94

Query Match 30.7%; Score 482; DB 17; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 1149
DB 500 CAGGATATTCCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGG 441
QY 1150 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 1209
DB 440 GGAATTAACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAGTTC 381
QY 1210 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 1269
DB 380 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCTGAGTCATCCAGAGACTCAGACA 321

```
QY 1270 CCGGCAATCCCACTGCTGACAGGAGACCCCTGACACTCTCTTCAAGACCTCTATTCCTTC 1329
DB 320 CCGGCAATCCCACTGCTGACAGGAGACCCCTGACACTCTCTTCAAGACCTCTATTCCTTC 261
QY 1330 CCGAGATGTTGAGATGTTGATCTCTCCAGCCCTGACCCCTGATGTTCTCTGATCAGG 1389
DB 260 CCGAGATGTTGAGATGTTGATCTCTCCAGCCCTGACCCCTGATGTTCTCTGATCAGG 201
QY 1390 GTCTGCTTCCCACTGAGGCTGACCGGCTCTCTGATGTTGAACCTGAGGAACAATTC 1449
DB 200 GTCTGCTTCCCACTGAGGCTGACCGGCTCTCTGATGTTGAACCTGAGGAACAATTC 141
QY 1450 CAAACTGTTCCAGGCGGGGTTGCTCTCATCTCTCCGAGCACTTTCATCTCAAGC 1509
DB 140 CAAACTGTTCCAGGCGGGGTTGCTCTCATCTCTCCGAGCACTTTCATCTCAAGC 81
QY 1510 TCAGGCGCCATCCCTTCTGACGCTCTGACCCCAATTTAGTCCGAGAAATAAAGTGA 1569
DB 80 TCAGGCGCCATCCCTTCTGACGCTCTGACCCCAATTTAGTCCGAGAAATAAAGTGA 21
QY 1570 AG 1571
DB 20 AG 19
```

RESULT 6
US-10-172-118-1411
Sequence 1411, Application US/10172118
Publication No. US20030224374A1

GENERAL INFORMATION:

APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Linsley, Peter
APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1411
LENGTH: 1260
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM 012427
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.4e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1090 CAGGATGATTCGGGGGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1149
DB 766 CAGGATGATTCGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 825
QY 1150 GGAAGATTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
DB 826 GGAAGATTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
QY 1210 ACCAAGTGTATCCAGGAACCATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1269
DB 886 ACCAAGTGTATCCAGGAACCATCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 945
QY 1270 CCGGCAATCCCACTGCTGACAGGAGACCCCTGACACTCTCTTCAAGACCTCTATTCCTTC 1329
DB 946 CCGGCAATCCCACTGCTGACAGGAGACCCCTGACACTCTCTTCAAGACCTCTATTCCTTC 1005
```

```
QY 1330 CCGAGATGTTGAGATGTTGATCTCTCCAGCCCTGACCCCTGATGTTCTCTGATCAGG 1389
DB 1006 CCGAGATGTTGAGATGTTGATCTCTCCAGCCCTGACCCCTGATGTTCTCTGATCAGG 1065
QY 1390 GTCTGCTTCCCACTGAGGCTGACCGGCTCTCTGATGTTGAACCTGAGGAACAATTC 1449
DB 1066 GTCTGCTTCCCACTGAGGCTGACCGGCTCTCTGATGTTGAACCTGAGGAACAATTC 1125
QY 1450 CAAACTGTTCCAGGCGGGGTTGCTCTCATCTCTCCGAGCACTTTCATCTCAAGC 1509
DB 1126 CAAACTGTTCCAGGCGGGGTTGCTCTCATCTCTCCGAGCACTTTCATCTCAAGC 1185
QY 1510 TCAGGCGCCATCCCTTCTGACGCTCTGACCCCAATTTAGTCCGAGAAATAAAGTGA 1569
DB 1186 TCAGGCGCCATCCCTTCTGACGCTCTGACCCCAATTTAGTCCGAGAAATAAAGTGA 1245
QY 1570 AG 1571
DB 1246 AG 1247
```

RESULT 7
US-10-295-027-515
Sequence 515, Application US/10295027
Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afari, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Bos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 515
LENGTH: 1260
TYPE: DNA
ORGANISM: Homo sapiens
US-10-295-027-515

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3.4e-132;

	Matches	482;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1090	CAGGGTGA	TTCTGGGGGGG	CGTGTG	CTGCAATGGCTCC	CTGCAGGAACTG	TGTCCTGG	1149		
Db	766	CAGGGTGA	TTCTGGGGGGG	CGTGTG	CTGCAATGGCTCC	CTGCAGGAACTG	TGTCCTGG	825		
QY	1150	GGAGTTA	CCCTTGGGCCCGG	CCCAAGACCGGGT	GTCTACGAACTCTG	CAAGTTC	1209			
Db	826	GGAGTTA	CCCTTGGGCCCGG	CCCAAGACCGGGT	GTCTACGAACTCTG	CAAGTTC	885			
QY	1210	ACCAAGTGA	CTCAGAAACCA	CTCAGAGCACTCTG	AGTCATCCAGACTC	GACACA	1268			
Db	886	ACCAAGTGA	CTCAGAAACCA	CTCAGAGCACTCTG	AGTCATCCAGACTC	GACACA	945			
QY	1270	CCGGCAT	CCCACTGCTGC	AGGAGCAGCCCTG	ACACTCTTTCAGACCTCAT	TCCTTC	1329			
Db	946	CCGGCAT	CCCACTGCTGC	AGGAGCAGCCCTG	ACACTCTTTCAGACCTCAT	TCCTTC	1009			
QY	1330	CCAGAA	TGTTGAATGTTCA	TCTCTCAAGCCCTG	ACCCGATGTCTCTG	AGACTCAGG	1388			
Db	1006	CCAGAA	TGTTGAATGTTCA	TCTCTCAAGCCCTG	ACCCGATGTCTCTG	AGACTCAGG	1065			
QY	1390	GTGTGCT	CCCCCACTTGGGG	GTGACGGTGTCTCTG	ATTGAACCGGGAA	CAATTC	1449			
Db	1066	GTGTGCT	CCCCCACTTGGGG	GTGACGGTGTCTCTG	ATTGAACCGGGAA	CAATTC	1125			
QY	1450	CAAAACTG	TCAGGGCGGGGG	TGGGTCTCAATCTCC	TGGGGCACTTTCATCTC	CAAGC	1509			
Db	1126	CAAAACTG	TCAGGGCGGGGG	TGGGTCTCAATCTCC	TGGGGCACTTTCATCTC	CAAGC	1185			
QY	1510	TCAGGG	CCCATCCCTTCTCTG	CAGCTTGACCCAAATTTAGTCC	CAAAATPAACTG	AGA	1568			
Db	1186	TCAGGG	CCCATCCCTTCTCTG	CAGCTTGACCCAAATTTAGTCC	CAAAATPAACTG	AGA	1245			
QY	1570	AG	1571							
Db	1246	AG	1247							

```

RESULT 8
US-10-173-999-53
Sequence 53, Application US/10173999
Publication No. US2004005563A1
GENERAL INFORMATION:
APPLICANT: Mack, David H.
APPLICANT: Gish, Kurt C.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
TITLE OF INVENTION: Cancer
FILE REFERENCE: 018501-002420US
CURRENT APPLICATION NUMBER: US/10/173.999
CURRENT FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: US 60/299,234
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: US 60/315,287
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/372,246
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 163
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 53
LENGTH: 1260
TYPE: DNA
ORGANISM: Homo sapiens
US-10-173-999-53

Query Match          30.7%   Score 482;   DB 17;   Length 1260;
Best Local Similarity 100.0%;   Pident No. 3; 4e-12;
Matches 482;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```

QY	1090	CAGGGTGAATTCGGGGGGCCCTGTAGTCTGCAATGGCTCCCGCAGGGACATGATCTCCGG	1149
Db	766	CAGGGTGAATTCGGGGGGCCCTGTAGTCTGCAATGGCTCCCGCAGGGACATGATCTCCGG	825
QY	1150	GGAAGTATTCCTGTGGCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAATTC	1209
Db	826	GGAAGTATTCCTGTGTGCCCCGGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAATTC	885
QY	1210	ACCAAGTGAATCCAGAAACCATCCAGGCCAACTCTGAAGTCATCCAGACTCAGACA	1265
Db	886	ACCAAGTGAATCCAGAAACCATCCAGGCCAACTCTGAAGTCATCCAGACTCAGACA	945
QY	1270	CCGGCATTCCCAACCTGTGTGCGAGGACAGCCCTGACATCTCTTTAGAACCTTCATTCCTTC	1329
Db	946	CCGGCATTCCCAACCTGTGTGCGAGGACAGCCCTGACATCTCTTTAGAACCTTCATTCCTTC	1005
QY	1330	CCAGAGATGTGAGAAATGTCATGTCCTCAGCCCTGACCCCATGTCTCCGSACTCAGG	1389
Db	1006	CCAGAGATGTGAGAAATGTCATGTCCTCAGCCCTGACCCCATGTCTCTCGSACTCAGG	1065
QY	1390	GTCGTCTTCCCCCAATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGAAACAATTC	1449
Db	1066	GTCGTCTTCCCCCAATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGAAACAATTC	1125
QY	1450	CAAAACTGTCCAGGGCCGGGGGTTCCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC	1509
Db	1126	CAAAACTGTCCAGGGCCGGGGGTTCCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC	1185
QY	1510	TCAGGGCCCAATCCCTTCTCTGCACTGTGACCCCAAAATTTAGTCCCAAGAAATAACTGAGA	1565
Db	1186	TCAGGGCCCAATCCCTTCTCTGCACTGTGACCCCAAAATTTAGTCCCAAGAAATAACTGAGA	1245
QY	1570	AG 1571	
Db	1246	AG 1247	

```

RESULT 9
US-10-342-887-1411
; Sequence 1411, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1411
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1411

Query Match 30.7%; Score 482; DB 17; Length 1260;
Best Local Similarity 100.0%; Pred. No. 3,4e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1090 CAGGGTATTCTGGGGGCGCTGTGCAATGGCTCCCTGACGGAGACTGTCTCTGG 1149

```

Db 766 CAGGATGATTCGGGGGGCTGTGTGATCTGCAATGAGCTCCCTGAGGGAAGCTGTGTCTCG 825
QY 1150 GGAGATTACCTCTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db 826 GGAAGATTACCTCTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 885
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGACACA 1269
Db 886 ACCAAGTGAATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGACACA 945
QY 1270 CCGGATCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTCAAGCTCTCAATCTCTTC 1329
Db 946 CCGGATCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTCAAGCTCTCAATCTCTTC 1005
QY 1330 CCAAGATGTTGAAGATGTTCACTCTCTGAGCCCTGACCCCAATGTCCTGAGACTCAG 1389
Db 1006 CCAAGATGTTGAAGATGTTCACTCTCTGAGCCCTGACCCCAATGTCCTGAGACTCAG 1065
QY 1390 GTCTGCTTCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTGAAGCTCTGAGAACTTC 1449
Db 1066 GTCTGCTTCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTGAAGCTCTGAGAACTTC 1125
QY 1450 CAAACTGTTCAGGCGGGGGTGTGTCTCAATCTCTCTGAGGCACTTTTCATCTCAAGC 1509
Db 1126 CAAACTGTTCAGGCGGGGGTGTGTCTCAATCTCTCTGAGGCACTTTTCATCTCAAGC 1185
QY 1510 TCAGGCGCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 1186 TCAGGCGCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1245
QY 1570 AG 1571
Db 1246 AG 1247

RESULT 10
US-09-739-907-52
Sequence 52, Application US/09739907
Patent No. US20010012889A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/739,907
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1381
TYPE: DNA
ORGANISM: Homo sapiens
US-09-739-907-52

Query Match 30.7%; Score 482; DB 9; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATGATTCGGGGGGCTGTGTGATCTGCAATGAGCTCCCTGAGGGAAGCTGTGTCTCG 1149
Db 852 CAGGATGATTCGGGGGGCTGTGTGATCTGCAATGAGCTCCCTGAGGGAAGCTGTGTCTCG 911
QY 1150 GGAAGATTACCTCTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 1209

Db 912 GGAAGATTACCTTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 971
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGACACA 1269
Db 972 ACCAAGTGAATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGACACA 1031
QY 1270 CCGGATCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTCAAGCTCTCAATCTCTTC 1329
Db 1032 CCGGATCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTCAAGCTCTCAATCTCTTC 1091
QY 1330 CCAAGATGTTGAAGATGTTCACTCTCTGAGCCCTGACCCCAATGTCCTGAGACTCAG 1389
Db 1092 CCAAGATGTTGAAGATGTTCACTCTCTGAGCCCTGACCCCAATGTCCTGAGACTCAG 1151
QY 1390 GTCTGCTTCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTGAAGCTCTGAGAACTTC 1449
Db 1152 GTCTGCTTCCCACTGTGTGAGGAGACAGCCCTGACACTCTCTTGAAGCTCTGAGAACTTC 1211
QY 1450 CAAACTGTTCAGGCGGGGGTGTGTCTCAATCTCTCTGAGGCACTTTTCATCTCAAGC 1509
Db 1212 CAAACTGTTCAGGCGGGGGTGTGTCTCAATCTCTCTGAGGCACTTTTCATCTCAAGC 1271
QY 1510 TCAGGCGCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 1272 TCAGGCGCCATCCCTTCTGTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1331
QY 1570 AG 1571
Db 1332 AG 1333

RESULT 11
US-09-938-671-52
Sequence 52, Application US/09938671
Publication No. US20040002066A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 36 Human Secreted Proteins
FILE REFERENCE: P2022P1
CURRENT APPLICATION NUMBER: US/09/938,671
PRIOR FILING DATE: 2001-08-27
PRIOR APPLICATION NUMBER: 09/348,457
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: 60/070,567
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,692
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,704
PRIOR FILING DATE: 1998-01-07
PRIOR APPLICATION NUMBER: 60/070,658
NUMBER OF SEQ ID NOS: 196
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 52
LENGTH: 1381
TYPE: DNA
ORGANISM: Homo sapiens
US-09-938-671-52

Query Match 30.7%; Score 482; DB 11; Length 1381;
Best Local Similarity 100.0%; Pred. No. 3.6e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATGATTCGGGGGGCTGTGTGATCTGCAATGAGCTCCCTGAGGGAAGCTGTGTCTCG 1149
Db 852 CAGGATGATTCGGGGGGCTGTGTGATCTGCAATGAGCTCCCTGAGGGAAGCTGTGTCTCG 911
QY 1150 GGAAGATTACCTTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db 912 GGAAGATTACCTTGTGTGCGCCGCCCAAGACAGCGGGTGTCTACAGAACTCTGCAAGTTC 971
QY 1210 ACCAAGTGAATCCAGGAAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGACACA 1269

Db	972	ACCAAGTGGATCTCAGAGAAACCATCCAGGGCAACTCTGAGTCATCTCCAGAGCTCAGACA	1031
QY	1270	CCGGCATCCCCACCTGCTGCAAGGGACAGCCCTGACATCTCTTTCAGACCTCATTTCTTC	1329
Db	1032	CCGGCATCCCCACCTGCTGCAAGGGACAGCCCTGACATCTCTTTCAGACCTCATTTCTTC	1091
QY	1330	CCAGGATGTTGAGAAATGTTCACTCTCTCAAGCCCTGACCCCAATGTCCTCGAGACTCAGG	1389
Db	1092	CCAGGATGTTGAGAAATGTTCACTCTCTCAAGCCCTGACCCCAATGTCCTCGAGACTCAGG	1151
QY	1390	GTCCTACTCCCCACCATTTGGGCTGACCGTGTCCTCTAGTTGAACCTCGGGAAACAATTC	1449
Db	1152	GTCCTACTCCCCACCATTTGGGCTGACCGTGTCCTCTAGTTGAACCTCGGGAAACAATTC	1211
QY	1450	CAAAACTGTCCAGGGCGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC	1509
Db	1212	CAAAACTGTCCAGGGCGGGGGGTTCGCTCAATCTCCCTGGGGCACTTTCATCTCTCAAGC	1271
QY	1510	TCAGGGCCCATCTCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA	1569
Db	1272	TCAGGGCCCATCTCTTCTCTGCAAGCTCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA	1331
QY	1570	AG	
Db	1332	AG 1333	

```

RESULT 12
US-09-739-907-53
; Sequence 53, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/739,907
; CURRENT FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-739-907-53

```

	Query Match	30.7%	Score 482	DB 9	Length 1439
	Best Local Similarity	100.0%	Pred. No. 3,66-112		
	Matches 482	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1090	CAGGATGATTTCTGGGGGGGCGCTGTGTCTGCATATGGCTCCCTGCAGAGGACTCTGTCTCTGG	1149		
Db	910	CAGGATGATTTCTGGGGGGGCGCTGTGTCTGCATATGGCTCCCTGCAGAGGACTCTGTCTCTGG	969		
QY	1150	GGAGATTACCTTGTGCCCGGCGCCAAACAGACCGGGTGTCTACACGAACTCTGCAGATTC	1209		
Db	970	GGAGATTACCTTGTGTGCCCGGCGCCAAACAGACCGGGTGTCTACACGAACTCTGCAGATTC	1029		
QY	1210	ACCAAGTGGATTCACAGAAACATCCAGGCGCAATCTCTGAGTCATCCCGAGACTAGACA	1259		
Db	1030	ACCAAGTGGATTCAGAAACATCCAGGCGCAATCTCTGAGTCATCCCGAGACTAGACA	1089		
QY	1270	CCGGCATCCCACTGTCTGCAGAGGACAGCCCTGACACTCTTTTCAGACCTCATTTCTTC	1329		

Accession	Sequence	Position
Db	CGGGATCCCACTGGTCGACAGGGACAGCCCTGACATCTCTTTAGACCCCTCATCTCTTC	1149
QY	CCAGAGATGTTGAGAAATGTTTCATCTCTTCAGACCCCTTGACCCCATGTCCTCTGACTCAGG	1389
Db	CCAGAGATGTTGAGAAATGTTTCATCTCTTCAGACCCCTTGACCCCATGTCCTCTGACTCAGG	1209
QY	GTCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCCGTGGACCAATTTTC	1449
Db	GTCTGCTTCCCCACATTTGGGCTGACCGGTGTCTCTAGTTGAAACCCGTGGACCAATTTTC	1269
QY	CAAAACGTCCAGAGGGCGGGGGTTCGGTCTCAATTCCTCGTGGGACATTTCATCTCTCAAGC	1509
Db	CAAAACGTCCAGAGGGCGGGGGTTCGGTCTCAATTCCTCGTGGGACATTTCATCTCTCAAGC	1329
QY	TCAGGGCCCATCCCTTCTCTGACAGCTGACCCCAAAATTTAGTCCACAGAAATTAACCTGAGA	1569
Db	TCAGGGCCCATCCCTTCTCTGACAGCTGACCCCAAAATTTAGTCCACAGAAATTAACCTGAGA	1389
QY	AG AG 1571	
Db	AG AG 1391	

```

US-09-938-671-53
RESULT 13
US-09-938-671-53
; Sequence 53, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022p1
; CURRENT APPLICATION NUMBER: US/09/938,671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,652
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 53
; LENGTH: 1439
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-53

```

	Query March	30.7%	Score 482	DB 11	Length 1439	
	Best Local Similarity	100.0%	Prod. No. 3.6e-132			
	Matches 482	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1090	CAGGGTATTCGTGGGGGCGCTGTGGTCTGCAGTGGCTCCGTGACAGGACTGTGTCTGG	1149			
Db	910	CAGGGTATTCGTGGGGGCGCTGTGGTCTGCAGTGGCTCCGTGACAGGACTGTGTCTGG	969			
QY	1150	GGAGATTACCTTGTGCCCCGGGCCACAGACGCGGGTGTCTACAGAACTCTGCAATTTC	1209			
Db	970	GGAGATTACCTTGTGCCCCGGGCCACAGACGCGGGTGTCTACAGAACTCTGCAATTTC	1029			
QY	1210	ACCAAGGAGATCCGAGGAAACCATCAGGCCCAATCCCTGAGATCATCCAGAGACTCAGACA	1269			
Db	1030	ACCAAGTGGATCCAGGAAACCATCAGGCCCAATCCCTGAGATCATCCAGAGACTCAGACA	1089			
QY	1270	CCGGCATTCGCCACCTGCTGACAGGACAGCCCTGACATCTCTTTACAGACCTCTCATCTCTTC	1329			
Db	1090	CCGGCATTCGCCACCTGCTGACAGGACAGCCCTGACATCTCTTTACAGACCTCTCATCTCTTC	1149			
QY	1330	CCAGAGATGTTGAGATGTTTCATCTTCCAGCCCTGATCCCATGTCATCTCTGGACTCAGG	1389			

Db 1150 CCAGAGATGTGAGATGTTCAATCTCTCCAGCCCTGACCCCAATGTCCTCGAGACTCAGG 1209
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1449
Db 1210 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1269
QY 1450 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1509
Db 1270 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1329
QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1569
Db 1330 TCAGGGCCCATCTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATGAGA 1389
QY 1570 AG 1571
Db 1390 AG 1391

RESULT 14

US-10-723-860-5032
; Sequence 5032, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Azile, Natasha
; APPLICANT: Gishburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPDS01
; CURRENT FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5032
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-5032

Query Match 30.7%; Score 482; DB 18; Length 1481;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGGAATCGTGTCTGG 1149
Db 894 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGGAATCGTGTCTGG 953
QY 1150 GGAAGATTAACCTTGGCCCGGCCCAACAGACCGGGGTGTACACGAACTCTGCAAGTTC 1209
Db 954 GGAAGATTAACCTTGGCCCGGCCCAACAGACCGGGGTGTACACGAACTCTGCAAGTTC 1013
QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1269
Db 1014 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1073
QY 1270 CCGGATATCCCACTGCTGCAAGGAGACCTCTGACACTCTCTTTCAGACCTCAATCTCTTC 1329
Db 1074 CCGGATATCCCACTGCTGCAAGGAGACCTCTGACACTCTCTTTCAGACCTCAATCTCTTC 1133
QY 1330 CCAGAGATGTGAGATGTTCAATCTCTCAGGCCCTGACCCCAATGTCCTCTGGAATCAGG 1389
Db 1134 CCAGAGATGTGAGATGTTCAATCTCTCAGGCCCTGACCCCAATGTCCTCTGGAATCAGG 1193
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1449
Db 1194 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1253
QY 1450 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1509
Db 1254 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1313

QY 1510 TCAGGGCCCATCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1569
Db 1314 TCAGGGCCCATCTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1373
QY 1570 AG 1571
Db 1374 AG 1375

RESULT 15

US-09-739-907-37
; Sequence 37, Application US/09739907
; Patent No. US20010012889A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT FILING DATE: 2000-12-20
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 1516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: n equals a,t,c, or c
US-09-739-907-37

Query Match 30.7%; Score 482; DB 9; Length 1516;
Best Local Similarity 100.0%; Pred. No. 3.7e-132;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGGAATCGTGTCTGG 1149
Db 977 CAGGGTGAATTCGGGGGGGCTGTGTCTGCAATGCTCCCTGAGGGAATCGTGTCTGG 1036
QY 1150 GGAAGATTAACCTTGGCCCGGCCCAACAGACCGGGGTGTACACGAACTCTGCAAGTTC 1209
Db 1037 GGAAGATTAACCTTGGCCCGGCCCAACAGACCGGGGTGTACACGAACTCTGCAAGTTC 1096
QY 1210 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1269
Db 1097 ACCAAGTGATCCAGGAAACCATCCAGGCCAACTCTGAGTCACTCCAGAGACTCAGACA 1156
QY 1270 CCGGATATCCCACTGCTGCAAGGAGACCTCTGACACTCTCTTTCAGACCTCAATCTCTTC 1329
Db 1157 CCGGATATCCCACTGCTGCAAGGAGACCTCTGACACTCTCTTTCAGACCTCAATCTCTTC 1216
QY 1330 CCAGAGATGTGAGATGTTCAATCTCTCAGGCCCTGACCCCAATGTCCTCTGGAATCAGG 1389
Db 1217 CCAGAGATGTGAGATGTTCAATCTCTCAGGCCCTGACCCCAATGTCCTCTGGAATCAGG 1276
QY 1390 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1449
Db 1277 GTCGTCTTCCCCACATTGGGCTGACCGGTCTCTAGTTGAACCTGGGAACAATTTC 1336
QY 1450 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1509
Db 1337 CAAAATGTCAGGGGGGGGGTTCAGTCTCAATCTCTCGGGCACTTTCATCTCAAGC 1396

Qy	1510	TCAGGGCCCATCCCTTCTCTGCAAGCTCTGA	CCCAATTAGTCCCGAATAAAGTGAGA	1569
Db	1397	TCAGGGCCCATCCCTTCTCTGCAAGCTCTGA	CCCAATTAGTCCCGAATAAAGTGAGA	1456
Qy	1570	AG	1571	
Db	1457	AG	1458	

Search completed: February 27, 2005, 04:52:32
job time : 850.613 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 20:26:59 ; Search time 4923.95 Seconds
(without alignments)
12144.524 Million cell updates/sec

Title: US-09-936-271b-13_COPY_10000_11570

Perfect score: 1571
1 agggaggtatgggggaattga.....cccaagaataactggaag 1571

Sequence:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Scoring table:

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	482	30.7	539	2	BF513278 UT-H-BM1-
C 2	482	30.7	586	7	W73140 zds5e11.81
C 3	475.8	30.3	809	4	BG680075 BG680075
C 4	455	29.0	478	1	AA862032 AA862032
C 5	443.8	28.2	802	4	BG682309 BG682309
C 6	440	28.0	953	2	BE745465 BE745465
C 7	431.8	27.5	666	4	BM842155 BM842155
C 8	417	26.5	645	1	AI002163 AI002163
C 9	383.6	24.4	677	2	BE898385 BE898385
C 10	381.8	24.3	453	7	NB0762 NB0762
C 11	355.8	22.6	363	2	AM105502 AM105502
C 12	348	22.2	590	4	BM840511 BM840511
C 13	333.4	21.2	577	4	BM838406 BM838406
C 14	322.8	20.5	430	7	W68361 W68361
C 15	317	20.2	388	7	W68496 W68496
C 16	291	18.5	533	2	BF514439 BF514439
C 17	291	18.5	533	4	BM840509 BM840509
C 18	278.8	17.7	523	4	BM841293 BM841293
C 19	277.2	17.6	320	2	AM380655 AM380655
C 20	276.6	17.6	411	2	AM801647 AM801647
C 21	273	17.4	526	4	BM830263 BM830263
C 22	270.4	17.2	526	4	BM841697 BM841697
C 23	245	15.6	498	7	W73168 W73168
C 24	229.8	14.6	420	6	CB298043 CB298043

25	229.2	14.6	422	4	BG398353
26	227.8	14.5	947	4	BG680848
27	227.6	14.5	479	2	AM576251
28	227.6	14.5	486	1	AI755214
29	227.6	14.5	726	5	BQ005946
30	227.6	14.5	726	7	CN395303
31	227.6	14.5	961	3	BU197099
32	227.6	14.5	4544	5	HSMB03689
33	226	14.4	359	2	AM500684
34	226	14.4	476	1	AI754567
35	226	14.4	481	1	AI754105
36	226	14.4	740	7	CN395308
37	226	14.4	977	6	CD516130
38	226	14.4	1092	4	BM480074
39	226	14.4	1547	3	CR621085
40	226	14.4	1547	3	CR615950
41	223.4	14.2	3341	3	CR749847
42	223	14.2	744	4	BG743962
43	222.4	14.2	715	2	BF574982
44	222	14.1	352	1	AI249688
45	221.2	14.1	649	4	BG393464

ALIGNMENTS

RESULT 1
BF513278/c
LOCUS
DEFINITION
US-H-BM1-amo-e-03-0-UI.81 NCI CGAP_Sub7 Homo sapiens cDNA clone
IMAGE:3070564 3', mRNA sequence.
ACCESSION
BF513278
VERSION
BF513278.1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 539)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
Tumor Gene Index
Unpublished (1997)
JOURNAL
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/ILMIL at:
www.bio.lml.gov/dbip/image/image.html
Seq primer: M13 Forward
POLVA=Yes.
FEATURES
source
Location/Qualifiers
1..539
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3070564"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP Sub7"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCI-CGAP Sub7
is a subcloned library derived from NCI CGAP Sub6. The
NCI CGAP Sub7 library had 12 million recombinants. A
single-stranded DNA preparation of NCI CGAP Sub6 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCI-CGAP_Kids pool 1 LLM
3334-3337, 3682-3683, 3798-3803 (IMAGE Clones)
1322376-1323911, 1456008-1456775, 1500552-1500855);
NCI CGAP Kids pool 1 LLM 3338-3342, 3722-3725, 3776-3778
(IMAGE Clones) 1323912-1325831, 1471368-1472803,

QY 1330 CCAAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTGAGCTCAG 1389
DB 261 CCAAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTGAGCTCAG 202
QY 1390 GTCTGCTTCCCAATGAGCTGACCCGCTCTCTGAGTTGAACCTTGGAACAATTC 1449
DB 201 GTCTGCTTCCCAATGAGCTGACCCGCTCTCTGAGTTGAACCTTGGAACAATTC 142
QY 1450 CAAATCTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGCACTTTCTCTCAGC 1509
DB 141 CAAATCTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGCACTTTCTCTCAGC 82
QY 1510 TCAGGGCCCATCTCTCTGACGCTGACCCCAATTTAGTCCCAAGAAATTAAGTGA 1569
DB 81 TCAGGGCCCATCTCTCTGACGCTGACCCCAATTTAGTCCCAAGAAATTAAGTGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 3
BG680075 809 bp mRNA linear EST 01-MAY-2001
LOCUS 602628224F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753372 5',
DEFINITION mRNA sequence.
ACCESSION BG680075 GI:13911472
VERSION BG680075.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 809)
AUTHORS NIH-NCI http://imgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM10612 row: 0 column: 05
High quality sequence stop: 707.
Location/Qualifiers
1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4753372"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1lb="NCI_CGAP_Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NCI;
Site 2: Sall; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 30.3%; Score 475.8; DB 4; Length 809;
Best Local Similarity 99.6%; Pired. No. 3.8e-64;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1084 CTCCTGAGGATGATTCCTGAGGGGCTGCTGCAATGCTCCCGAGGAGACTGCTG 1143
DB 219 CTCCTGAGGATGATTCCTGAGGGGCTGCTGCAATGCTCCCGAGGAGACTGCTG 278
QY 1144 TCCTGGGAGATTACCTTGTGCGCGGCCCAAGACCGGGTCTTACAGAACTCTGCG 1203

DB 279 TCCTGGGAGATTACCTTGTGCGCGGCCCAAGACCGGGTCTTACAGAACTCTGCG 338
QY 1204 AAGTTACCAATGATTCAGAGAAACATCCAGGCCCAATCTCTGATCATCTCCAGAGATTC 1263
DB 339 AAGTTACCAATGATTCAGAGAAACATCCAGGCCCAATCTCTGATCATCTCCAGAGATTC 398
QY 1264 AGCAGACCGGATTCCTGCTGAGGAGAGAGCCCTGACACTCTCTTCAAGACCTCAT 1323
DB 399 AGCAGACCGGATTCCTGCTGAGGAGAGAGCCCTGACACTCTCTTCAAGACCTCAT 458
QY 1324 TCCTTCCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTCTGGA 1383
DB 459 TCCTTCCAGAGATGTTGAAATGTTGATCTCTGACGCCCTGACCCCAATGCTCTCTGGA 518
QY 1384 CTCAGGCTCTGCTTCCCAATGAGCTGACCCGCTCTCTGAGTTGAACCTTGGAAC 1443
DB 519 CTCAGGCTCTGCTTCCCAATGAGCTGACCCGCTCTCTGAGTTGAACCTTGGAAC 578
QY 1444 AATTTCCAAACTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGCACTTTCTATCC 1503
DB 579 AATTTCCAAACTGTCAGAGGCGGGGGTTCGCTCTCAATCTCTGAGGCACTTTCTATCC 638
QY 1504 TCAAGCTCAGAGGCCCATCCCTTCTCTGACGCTTGAACCCCAATTTAGTCCCAAGAAATTA 1562
DB 639 TCAAGCTCAGAGGCCCATCCCTTCTCTGACGCTTGAACCCCAATTTAGTCCCAAGAAATTA 697

RESULT 4
AA862032 478 bp mRNA linear EST 13-APR-1999
LOCUS 0146609.s1 NCI_CGAP_HN3 Homo sapiens cDNA clone IMAGE:1485736 3',
DEFINITION similar to SW_SCE_HUMAN P49862 STRATUM CORNEUM CHROMOTRYPIC ENZYME
PRECURSOR ;, mRNA sequence.
ACCESSION AA862032 GI:2954511
VERSION AA862032.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 478)
AUTHORS NCI/NIH-CDAP http://www.ncbi.nlm.nih.gov/ncicdap.
TITLE National Cancer Institute / National Institute of Dental Research,
Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: John Emley, M.D.
CDNA Library Preparation: Stratagene, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CDAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert length: 1647 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 233.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1485736"
/tissue_type="squamous cell carcinoma from base of tongue"
/lab_host="SOLR (kanamycin resistant)"
/clone_1lb="NCI_CGAP_HN3"
/note="Organ: tongue; Vector: Bluescript SK-; Site: 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Average insert size 1.0 kb. 5' adaptor sequence:
5' GAATTCGACGAG 3' 3' adaptor sequence: 5'
(GA)10ACTAGTCTCAGATTCTTTTCTTTTCTTTT 3'."

ORIGIN

Query Match 29.0%; Score 455; DB 1; Length 478;
Best Local Similarity 99.8%; Pred. No. 7.5e-61;
Matches 466; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1105 GGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGGGAGATTACCTTGT 1164
DB 478 GGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTGTCTCTGGGAGATTACCTTGT 420
QY 1165 GCGCGGCGCAAGACCGGGGTGTCTACAGAACTCTGCAAGTTCACCAAGTGGATCCAG 1224
DB 419 GCGCGGCGCAAGACCGGGGTGTCTACAGAACTCTGCAAGTTCACCAAGTGGATCCAG 360
QY 1225 GAAACCATCAAGCCCACTCTGAGTCAATCCAGAGCTAGCAACACCGGATCCCACT 1284
DB 359 GAAACCATCAAGCCCACTCTGAGTCAATCCAGAGCTAGCAACACCGGATCCCACT 300
QY 1285 GCTGCAAGGACAGCCCTGACATCTCTTTGACACCTTCATTTCTCCAGAGATTTGAGA 1344
DB 299 GCTGCAAGGACAGCCCTGACATCTCTTTGACACCTTCATTTCTCCAGAGATTTGAGA 240
QY 1345 ATGTTCACTCTCCAGACCCCTGACCCCATGTCTCTGAGCTCAGGGTCTGTTCCCCAC 1404
DB 239 ATGTTCACTCTCCAGACCCCTGACCCCATGTCTCTGAGCTCAGGGTCTGTTCCCCAC 180
QY 1405 ATGAGGCTGACCGTGTCTCTGATGAAACCTGAGGACAAATTTCCAAAATGTCCAGG 1464
DB 179 ATGAGGCTGACCGTGTCTCTGATGAAACCTGAGGACAAATTTCCAAAATGTCCAGG 120
QY 1465 CCGGGGTTCCGTCTCAATCTCCCTGGGGCACTTCACTCTCAAGCTCAGGGCCCATCCCT 1524
DB 119 CCGGGGTTCCGTCTCAATCTCCCTGGGGCACTTCACTCTCAAGCTCAGGGCCCATCCCT 60
QY 1525 TCTCTGACCTGACCCCAATTTAGTCCAGAAATTAACCTGAGAG 1571
DB 59 TCTCTGACCTGACCCCAATTTAGTCCAGAAATTAACCTGAGAG 13

RESULT 5
BG682309 802 bp mRNA 1linear EST 01-MAY-2001
LOCUS 60269926F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',
DEFINITION mRNA Sequence.
ACCESSION BG682309
VERSION BG682309.1 GI:13913706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM10615 row: d column: 08
Location/Qualifiers
1. 802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754263"
/issue_type="equamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NCI_CGAP_Skn4"

/note="Organ: skin; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

Query Match 28.2%; Score 443.8; DB 4; Length 802;
Best Local Similarity 99.3%; Pred. No. 3.5e-59;
Matches 456; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1090 CAGGGTATTCTGGGAGGCTGTGTCTGCAATGCTCTCCAGAGACTCTGTCTCG 1149
DB 333 CAGGGTATTCTGGGAGGCTGTGTCTGCAATGCTCTCCAGAGACTCTGTCTCG 392
QY 1150 GGAAGATTACCTTGCCCGGCGCAAGACCGGGGTCTACAGAACTCTGCAAGTTC 1209
DB 393 GGAAGATTACCTTGCCCGGCGCAAGACCGGGGTCTACAGAACTCTGCAAGTTC 452
QY 1210 ACCAAGTGATCCAGAAACCAATCCAGGCCAATCTCTGATCATCCAGAGCTCAGACA 1269
DB 453 ACCAAGTGATCCAGAAACCAATCCAGGCCAATCTCTGATCATCCAGAGCTCAGACA 512
QY 1270 CCGGATCCCACTGCTGCAAGGACAGCCCTGACATCTCTTTCAGACCTCATTCCTTC 1329
DB 513 CCGGATCCCACTGCTGCAAGGACAGCCCTGACATCTCTTTCAGACCTCATTCCTTC 572
QY 1330 CCAGAGATTGAGAAATGTAATCTCTCAGCCCGGACCCCATGCTCTGAGACTCAGG 1389
DB 573 CCAGAGATTGAGAAATGTAATCTCTCAGCCCGGACCCCATGCTCTGAGACTCAGG 632
QY 1390 GTCTGCTTCCCCACA-TTGGGCTGACCGTGTCTCTGATGAAACCTGAGAACATTT 1448
DB 633 GTCTGCTTCCCCACAATTTGGGCTGACCGTGTCTCTGATGAAACCTGAGAACATTT 692
QY 1449 CCAAAACTGTCCAGGGCGGGGGTTCGCTCAATCTCCCTGGGACATTTTACTCTCAAG 1508
DB 693 CCAAAACTGTCCAGGGCGGGGGTTCGCTCAATCTCCCTGGGACATTTTACTCTCAAG 752
QY 1509 CTCAGGGCCCAATCCCTTCTCTGACGCTGACCCCAATT 1547
DB 753 CTCAGGGCCCAATCCCTTCTCTGACGCTGACCCCAATT 791

RESULT 6
BE745465 953 bp mRNA 1linear EST 15-SEP-2000
LOCUS 601579834F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928640 5',
DEFINITION mRNA Sequence.
ACCESSION BE745465
VERSION BE745465.1 GI:10159457
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM760 row: c column: 09
Location/Qualifiers
1. 953
/organism="Homo sapiens"
/mol_type="mRNA"

ORIGIN

Query Match 28.0%; Score 440; DB 2; Length 953;
Best Local Similarity 99.6%; Pred. No. 1.3e-56;
Matches 462; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

/db xref="taxon:9606"
/clone="IMAGE:392864.0"
/issue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1ib="NH_MGC_9"
/note="Organ: ovary; Vector: pOTB1; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GCGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

1090 CAGGGTATTCCTGGGGGCGCTGTGCTGCAATGCTCCCTGACAGGACTCGTCTCTGG 1149
|||
229 CAGGGTATTCCTGGGGGCGCTGTGCTGCAATGCTCCCTGACAGGACTCGTCTCTGG 288
|||
1150 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
|||
289 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 348
|||
1210 ACCAAGTATTCAGAGAAACATTCAGAGCAATCTCTGAGTCAATCCAGACTCAGACA 1269
|||
349 ACCAAGTATTCAGAGAAACATTCAGAGCAATCTCTGAGTCAATCCAGACTCAGACA 408
|||
1270 CCGGCAATCCGACCTGCTGAGGAGGAGCCCTGCACTCCTTTCAGACCTCAATCTCTG 1329
|||
409 CCGGCAATCCGACCTGCTGAGGAGGAGCCCTGCACTCCTTTCAGACCTCAATCTCTG 468
|||
1330 CCAGAGATTTGGAATGTTTCAATCTCTGAGGAGGAGCCCTGCACTCCTTTCAGACCT 1389
|||
469 CCAGAGATTTGGAATGTTTCAATCTCTGAGGAGGAGCCCTGCACTCCTTTCAGACCT 528
|||
1390 GTCTGCTTCCCTCCACATGAGGCTGAGCCGTCTCTTCTGAGTGAACCTCTGGAACTTTC 1449
|||
529 GTCTGCTTCCCTCCACATGAGGCTGAGCCGTCTCTTCTGAGTGAACCTCTGGAACTTTC 587
|||
1450 CAAACTGTCCAGAGGAGGAGGAGTGTGCTCAATCTCCTGGGACCTTCAATCTCAAGC 1509
|||
588 CAAACTGTCCAGAGGAGGAGGAGTGTGCTCAATCTCCTGGGACCTTCAATCTCAAGC 647
|||
1510 TCAGGGCCCATCCCTCTCTGACGCTGACCCCAATTTAGTCC 1553
|||
648 TCAGGGCCCATCCCTCTCTGACG-TCTGACCCCAATTTAGTCC 690
|||

RESULT 7
BM842155 666 bp mRNA linear EST 06-MAR-2002
LOCUS K-E870119498 S12SNU216 Homo sapiens cDNA clone S12SNU216-56-H11 5',
DEFINITION mRNA sequence.
ACCESSION BM842155
VERSION BM842155.1 GI:19198564
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 666)
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Institute of Bioscience & Biotechnology

52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 56 row: H column: 11
High quality sequence stop: 666.
Location/Qualifiers
1. 666
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-56-H11"
/sex="F"
/issue_type="lymph node"
/cell_type="epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"
/note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then deacapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dT-selected mRNA by
priming with dT-tailed vector. The dT-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."

ORIGIN

Query Match 27.5%; Score 431.8; DB 4; Length 666;
Best Local Similarity 99.5%; Pred. No. 2.7e-57;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1090 CAGGGTATTCCTGGGGGCGCTGTGCTGCAATGCTCCCTGACAGGACTCGTCTCTGG 1149
|||
232 CAGGGTATTCCTGGGGGCGCTGTGCTGCAATGCTCCCTGACAGGACTCGTCTCTGG 291
|||
1150 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1209
|||
292 GGAATTAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 351
|||
1210 ACCAAGTATTCAGAGAAACATTCAGAGCAATCTCTGAGTCAATCCAGACTCAGACA 1269
|||
352 ACCAAGTATTCAGAGAAACATTCAGAGCAATCTCTGAGTCAATCCAGACTCAGACA 411
|||
1270 CCGGCAATCCGACCTGCTGAGGAGGAGCCCTGCACTCCTTTCAGACCTCAATCTCTG 1329
|||
412 CCGGCAATCCGACCTGCTGAGGAGGAGCCCTGCACTCCTTTCAGACCTCAATCTCTG 471
|||
1330 CCAGAGATTTGGAATGTTTCAATCTCTGAGGAGGAGCCCTGCACTCCTTTCAGACCT 1389
|||
472 CCAGAGATTTGGAATGTTTCAATCTCTGAGGAGGAGCCCTGCACTCCTTTCAGACCT 531
|||
1390 GTCTGCTTCCCTCCACATGAGGCTGAGCCGTCTCTGAGTGAACCTCTGGAACTTTC 1449
|||
532 GTCTGCTTCCCTCCACATGAGGCTGAGCCGTCTCTGAGTGAACCTCTGGAACTTTC 591
|||
1450 CAAACTGTCCAGAGGAGGAGGAGTGTGCTCAATCTCCTGGGACCTTCAATCTCAAGC 1509
|||
592 CAAACTGTCCAGAGGAGGAGGAGTGTGCTCAATCTCCTGGGACCTTCAATCTCAAGC 651
|||
1510 TCAGGGCCCATCCCT 1524
|||
652 TCAGGGCCCATCCCT 666
|||

```
RESULT 8
AI002163/c 645 bp mRNA linear EST 27-AUG-1998
LOCUS ocl4h05.sl Soares testis NHT Homo sapiens cDNA clone IMAGE:1619481
DEFINITION 3' similar to TR:061955 Q61955 SERINE PROTEASE INHIBITOR 5'; mRNA
sequence.
ACCESSION AI002163 GI:3202497
VERSION AI002163
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 645)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/bdrr/image/image.html
Insert Length: 1042 Std Error: 0.00
Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 303.
FEATURES
source
1. 645
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1619481"
/sex="male"
/lab_host="DH10B"
/clone_lib="Soares testis NHT"
/notes="Vector: pT7AD-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo (dT)
primer [5'
TGTACCATCTGAAGTGGAGGCGGCCCAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Col5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
```

```
ORIGIN
Query Match 26.5%; Score 417; DB 1; Length 645;
Best Local Similarity 94.9%; Pred. No. 5.3e-55;
Matches 464; Conservative 0; Mismatches 20; Indels 5; Gaps 3;
```

```
1084 CTCCTGACAGAGTATCTGGGGGCTGTGATGTGCAATGGCTCCGACAGGACTCGTG 1143
DB CTCCTGACAGAGTATCTGGGGGCTGTGATGTGCAATGGCTCCGACAGGACTCGTG- 441
1144 TCTCTGGGAGATTACCTTTGTGCGCCGCCAACAGACGGGGTCTTACAGCAACTCTTGC 1203
DB 440 --CTGTGGAGATTACCTTTGT-CCCGGCCCAACAGACGGGGTCTTACAGCAACTCTTGC 384
1204 AAGTTCACCAAGTGTTCACAGAAACATTCAGGCCCAATCTCTGATCATCCAGAGATC 1263
DB 383 AAGTTCACCAAGTGTTCACAGAAACATTCAGGCCCAATCTCTGATCATCCAGAGATC 324
1264 AGCAGACCGGCATCCCACTGCTGACAGGAGACGCCCTGACACTCTTTTCAAGACCTCAT 1323
DB 323 AGCAGACCGGCATCCCACTGCTGACAGGAGACGCCCTTACACTCTTTTCAAGACCTCAT 264
1324 TCTTCCACAGAGATTGTGAATGTTCATCTCCAGCCCTGACCCCATGTCTCTGGA 1383
```

```
DB 263 TCTTCCACAGAGATTGTGAATGTTCATCTCCAGCCCTGACCCCATGTCTCTGGA 204
1384 CTCAGAGTCTGTTCCCTCCCAATTTGGGCTGACCGGTCTCTAGTTGAACCTGGGAGC 1443
DB 203 CTCAGAGTCTGTTCCCTCCCAATTTGGGCTGACCGGTCTCTAGTTGAACCTGGGAGC 144
1444 AATTTCCAAAATCTGTCCAGAGGCGGGGGTTCCTCTCAATCTCTCTGGGCACTTTCATCC 1503
DB 143 AATTTCCAAAATCTGTCCAGAGGCGGGGGTTCCTCTCAATCTCTCTGGGCACTTTCATCC 84
1504 TCAAGTCAAGGCGCCCATCTCTCTGACGCTGACCCCAATTAGT-CCCAAGAAATA 1562
DB 83 TCAAGTCAAGGCGCCCATCTCTCTGACGCTGACCCCAATTAGTCCCAAGAAATA 24
1563 ACTGAGAG 1571
DB 23 ACTGAGAG 15
```

```
RESULT 9
BE898385 677 bp mRNA linear EST 29-SEP-2000
LOCUS 601681219f1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951365 5',
DEFINITION mRNA sequence.
ACCESSION BE898385
VERSION BE898385.1 GI:10364809
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at: image.lnl.gov
Place: LINC819 row: f column: 06
High quality sequence stop: 677.
FEATURES
source
1. 677
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951365"
/risue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_9"
/notes="Organ: Ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming the following 5'
cloned into EcoRI/XhoI sites using the directionally
adaptor: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```
Query Match 24.4%; Score 383.6; DB 2; Length 677;
Best Local Similarity 98.5%; Pred. No. 7.8e-50;
Matches 397; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
```

```
1090 CAGGGTGAATTCTGGGGGCTGTGATGTGCAATGGCTCCGACAGGACTCGTCTCGG 1149
DB 276 CAGGGTGAATTCTGGGGGCTGTGATGTGCAATGGCTCCGACAGGACTCGTCTCGG 335
```

QY 1150 GGAGATTACCTTGTGCGCCGCCAAGACAGACCGGGGTCTTACAGAACTCTGCAATTTC 1209
 |||||
 Db 336 GGAGATTACCTTGTGCGCCGCCAAGACAGACCGGGGTCTTACAGAACTCTGCAATTTC 395
 |||||
 QY 1210 ACCAAGTGATTCAGAAACCATTCAGGCCCACTCTGAGTCATCCAGAGACTAGACA 1269
 |||||
 Db 396 ACCAAGTGATTCAGAAACCATTCAGGCCCACTCTGAGTCATCCAGAGACTAGACA 455
 |||||
 QY 1270 CCGGATCCCTCCTGCTGAGAGGACGCGCTGACATCTCTTTCAGACCTTCATCTCTTC 1329
 |||||
 Db 456 CCGGATCCCTCCTGCTGAGAGGACGCGCTGACATCTCTTTCAGACCTTCATCTCTTC 515
 |||||
 QY 1330 CCAGAGATGTTGAAATGTTTCACTCTCCAGACCCCTGACCATGCTCTCTGAGACTGAG 1389
 |||||
 Db 516 CCAGAGATGTTGAAATGTTTCACTCTCCAGACCCCTGACCATGCTCTCTGAGACTGAG 575
 |||||
 QY 1390 GTCTGCTTCCCAATGAGGCTGACCGGTCTCTCTAGTTGAAACCTTGGAACAATTTC 1449
 |||||
 Db 576 GNTCTGTTCCCAATGAGGCTGACCGGTCTCTCTAGTTGAAACCTTGGAACAATTTC 635
 |||||
 QY 1450 CAAACTCTCCAGGCGCGGGGTTGCTCTCAATCTCCCTGGGG 1492
 |||||
 Db 636 CAAACTCTCCAGGCGCGGGGTTGCTCTCAATCTCCCTGGGG 677
 |||||

RESULT 10
 N80762/c 453 bp mRNA linear EST 29-MAR-1996
 LOCUS 298106.g1 Soares fetal lung_NbH19w Homo sapiens cDNA clone
 DEFINITION IMAGE:300611.3, mRNA sequence.
 ACCESSION N80762
 VERSION N80762.1 GI:1243463
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
 TITLE The WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LINT; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: m13 -40 forward
 High quality sequence stop: 319.
 Location/Qualifiers
 1..453
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1245535"
 /db_xref="taxon:9606"
 /clone="IMAGE:300611"
 /dev_page="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_1b="Soares fetal lung NbH19w"
 /note="Organ: lung; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAATGGAGCGCGCAATTTTCTTTTCTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73D vector (Pharmacia). Library went through one round of

ORIGIN

normalization to a Cot = 5. Library constructed by Bento Soares and M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NbH19w."

Query Match 24.3%; Score 381.8; DB 7; Length 453;
 Best Local Similarity 96.6%; Pred. No. 1.7e-49;
 Matches 422; Conservative 0; Mismatches 9; Indels 6; Gaps 3;

QY 1141 GTGCTCTGGAGATTAACCTTGTGCGCCGCCAAGACGAGAGGAGTCTTACAGAA 1196
 |||||
 Db 441 GTGCTCTGGAGATTAACCTTGTGCGCCGCCAAGAGAGGAGTCTTACAGAA 382
 |||||
 QY 1197 CCTCTGCAAGTTACC-AGTGAATCCAGAAACCATTCAGGCCAATCTCTGAGTATCC 1255
 |||||
 Db 381 CTTCTCAAGTTACCAAGATGATCAGAGAAACATTCAGGCCAATCTCTGAGTATCC 322
 |||||
 QY 1256 CAGGACTCAGACACCGGCAATCCCACTGAC-TCAGAGGACAGCCCTGACACTCTTTC 1314
 |||||
 Db 321 CAGGACTCAGACACCGGCAATCCCACTGAC-TCAGAGGACAGCCCTGACACTCTTTC 262
 |||||
 QY 1315 GACCTCATTCCTTCCAGAGATGTTGAGAAATGTCATCTCTCAGCCCTGACCCCATG 1374
 |||||
 Db 261 GACCTCATTCCTTCCAGAGATGTTGAGAAATGTCATCTCTCAGCCCTGACCCCATG 202
 |||||
 QY 1375 TCTCTGAGACTCAGGCTGCTTCCCTCCACATTTGGCTGACCTGCTCTCTAGTTGAC 1434
 |||||
 Db 201 TCTCTGAGACTCAGGCTGCTTCCCTCCACATTTGGCTGACCTGCTCTCTAGTTGAC 142
 |||||
 QY 1435 CTTGGGAAACAATTTCCAAAACCTGTCAGAGGCGGGGTTGCTCATCTCTCTGAGGAC 1494
 |||||
 Db 141 CTTGGGAAACAATTTCCAAAACCTGTCAGAGGCGGGGTTGCTCATCTCTCTGAGGAC 82
 |||||
 QY 1495 CTTTCATCTCAAGCTCAGGCGCCCATCTCTCTGACGCTGACCCAAATTAATGCC 1554
 |||||
 Db 81 CTTTCATCTCAAGCTCAGGCGCCCATCTCTCTGACGCTGACCCAAATTAATGCC 22
 |||||
 QY 1555 AGAATTAACCTGAGAG 1571
 |||||
 Db 21 AGAATTAACCTGAGAG 5
 |||||

RESULT 11
 AM105502/c 363 bp mRNA linear EST 20-OCT-1999
 LOCUS x053906.x1 NCI_CGAP_Ov23 Homo sapiens cDNA clone IMAGE:2597530.3,
 DEFINITION mRNA sequence.
 ACCESSION AM105502
 VERSION AM105502.1 GI:6076237
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmerit-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at:
 www-bio.llnl.gov/btrp/image/image.html

Possible reversed clone: polyT not found
 Seq primer: -40UP from Gbco

FEATURES High quality sequence stop: 362.
Location/Qualifiers
source 1..363
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2597530"
/issue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCL CGAP OY23"
/note="Organ: OTary; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"

ORIGIN

Query Match 22.6%; Score 355.8; DB 2; Length 363;
Best Local Similarity 99.4%; Pred. No. 1.9e-45;
Matches 357; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1213 AAGTGATCCAGAAACCAATCCAGGCACTCTGTGATCATCCAGACTCAGACACCG 1272
DB 363 AAGTGATCCAGAAACCAATCCAGGCACTCTGTGATCATCCAGACTCAGACACCG 304
QY 1273 GCATCCCACTGCTCAGGAGACAGCCCTGACACTCTTTCAGACCTCATTCCTCCCA 1332
DB 303 GCATCCCACTGCTCAGGAGACAGCCCTGACACTCTTTCAGACCTCATTCCTCCCA 244
QY 1333 GAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCAATGTTCTCTGAGACTCAGGGTC 1392
DB 243 GAGATGTGAGAAATGTTCACTCTCCAGCCCTGACCCCAATGTTCTCTGAGACTCAGGGTC 184
QY 1393 TGCTTCCCACTGAGGAGTGTGCTCAATCTCCCTGAGGCACTTTCATCTCAAGCTCA 1452
DB 183 TGCTTCCCACTGAGGAGTGTGCTCAATCTCCCTGAGGCACTTTCATCTCAAGCTCA 124
QY 1453 AACTGTCCAGGAGGAGGAGTGTGCTCAATCTCCCTGAGGCACTTTCATCTCAAGCTCA 1512
DB 123 AACTGTCCAGGAGGAGGAGTGTGCTCAATCTCCCTGAGGCACTTTCATCTCAAGCTCA 64
QY 1513 GGGCCCATCTCTCTGCACTCTGACCCCAATTTAGTCCAGAAATAAAGTGAAG 1571
DB 63 GGGCCCATCTCTCTGCACTCTGACCCCAATTTAGTCCAGAAATAAAGTGAAG 5

RESULT 12
BM840511 590 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0117576 S12SNU216 Homo sapiens cDNA clone S12SNU216-39-G08 5',
DEFINITION mRNA sequence.
ACCESSION BM840511
VERSION BM840511 GI:19196920
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 590)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: Yongsung@mail.kribb.re.kr
Plate: 39 row: G column: 08

FEATURES High quality sequence stop: 590.
Location/Qualifiers
source 1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-39-G08"
/sex="F"
/issue_type="Lymph node"
/cell_type="Epithelial"
/cell_line="SNU-216"
/lab_host="Top10F"
/clone_lib="S12SNU216"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deapped with tabacco acid pyrophosphatase (TAP). The deapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoR I which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 22.2%; Score 348; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.7e-44;
Matches 348; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGCTGTGTGTCGAATGCTCTCCAGAGGATGCTGTCTTG 1149
DB 243 CAGGGTATCTGGGGGCTGTGTGTCGAATGCTCTCCAGAGGATGCTGTCTTG 302
QY 1150 GGAAGTTACCTTGCCCGCCGCAAGACCGGGGNGTCAACGAACTCTGCAAGTTC 1209
DB 303 GGAAGTTACCTTGCCCGCCGCAAGACCGGGGNGTCAACGAACTCTGCAAGTTC 362
QY 1210 ACCAAGTGATCCAGAAACCAATCCAGGCACTCTGATGATCCAGAGTCAACACA 1269
DB 363 ACCAAGTGATCCAGAAACCAATCCAGGCACTCTGATGATCCAGAGTCAACACA 422
QY 1270 CCGGCATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTCAAGACCTTATCTTC 1329
DB 423 CCGGCATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTCAAGACCTTATCTTC 482
QY 1330 CCAGAGATGTGAGATGTTCACTCTCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 1389
DB 483 CCAGAGATGTGAGATGTTCACTCTCAGCCCTGACCCCAATGCTCTCTGAGTCAAG 542
QY 1390 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTTATGTTGAACCT 1437
DB 543 GTCTGCTTCCCACTGAGGCTGACCGTGTCTCTTATGTTGAACCT 590

RESULT 13
BM838406 577 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5',
DEFINITION mRNA sequence.
ACCESSION BM838406
VERSION BM838406.1 GI:19194815
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 577)

AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Beon-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongseung@mail.krdb.re.kr
 Plate: 36 row: A column: 03
 High quality sequence stop: 577.

FEATURES**source**

1. 577
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="S12SN0216-36-A03"
 /sex="F"
 /tissue_type="Lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10P"
 /clone_11b="S12SN0216"
 /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 21.2%; Score 333.4; DB 4; Length 577;
Best Local Similarity 99.7%; Pred. No. 4.9e-42;
Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1090 CAGGGGATCTGGGGGGGCGTGTGCAATGGCTCCGTCAGAGGAGCTGGTCTCTG 1149
 |||||
 243 CAGGGGATCTGGGGGGGCGTGTGCAATGGCTCCGTCAGAGGAGCTGGTCTCTG 302
 |||||
 1150 GGAGATTACCTTGTGCGCCGCGCCACAGACCGGGTGTCTACAGAACTCTGCAATTC 1209
 |||||
 303 GGAGATTACCTTGTGCGCGCGCCACAGACCGGGTGTCTACAGAACTCTGCAATTC 362
 |||||
 1210 ACCAAGTGAATCCAGAAACCATTCAGGCCAATCTGAGTCAATCCAGAGACTCAGACA 1269
 |||||
 363 ACCAAGTGAATCCAGAAACCATTCAGGCCAATCTGAGTCAATCCAGAGACTCAGACA 422
 |||||
 1270 CCGGCAATCCCACTGTGTGAGGGAGAGCCCTGACATCTCTTCAGACCCCTCAATTCCTTC 1329
 |||||
 423 CCGGCAATCCCACTGTGTGAGGGAGAGCCCTGACATCTCTTCAGACCCCTCAATTCCTTC 482
 |||||
 1330 CCAGAGATGTGAGAAATGTTCAATCTCTCAGACCCCTGACCCCAATGTCTCTGAGACTGAG 1389
 |||||
 483 CCAGAGATGTGAGAAATGTTCAATCTCTCAGACCCCTGACCCCAATGTCTCTGAGACTGAG 542
 |||||
 1390 GTCTGCTTCCCAATTTGGGCTGAGCCGTGTCTCT 1424
 |||||
 543 GTCTGCTTCCCAATTTGGGCTGAGCCGTGTCTCT 577
 |||||

RESULT 14
LOCUS W68361/C
DEFINITION 430 bp mRNA linear EST 15-OCT-1996
 2d34n08.s1 Soares fetal heart NBH19W Homo sapiens cDNA clone
IMAGE:342591 3', mRNA sequence.

ACCESSION W68361
VERSION W68361.1 GI:1377251
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE**AUTHORS**

Hillier, L., Clark, N., Dubuque, T., Eliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Matra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

This clone is available royalty-free through LNL; contact the IMAG Consortium (info@image.lnl.gov) for further information.
 Insert Length: 400 Std Error: 0.00
 Seq primer: mob.RECA-ET

High quality sequence stop: 286.
 Location/Qualifiers

FEATURES**source**

1. 430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1267966"
 /db_xref="taxon:9606"
 /clone="IMAGE:342591"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_11b="Soares fetal heart NBH19W"
 /note="Organ: heart; Vector: pT773D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer (5' TGTTCACATCTGAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed from the M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBH19W."

ORIGIN

Query Match 20.5%; Score 322.8; DB 7; Length 430;
Best Local Similarity 98.9%; Pred. No. 2.4e-40;
Matches 346; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

1224 GGAACCATTCAGGGGCAATCTGAGTCAAT-CCAGAGATCAGACACGGGCAATCCCAAC 1282
 |||||
 359 GAAACCATTCAGGGGCAATCTGAGTCAATCCAGAGATCAGACACGGGCAATCCCAAC 300
 |||||
 1283 CTGCTCAGGAGACAGGCTGACACTCTTTCAGACGACCTATTCCTTCCAGAGATGTTG 1341
 |||||
 299 CTGCTCAGGAGACAGGCTGACACTCTTTCAGAGACGACCTATTCCTTCCAGAGATGTTG 240
 |||||
 1342 AGAATGTTCAATCTCTCAGAGCCCTGACCCCAATGTCTCTGAGACTCAGGGTCTGCTTCCC 1401
 |||||
 239 AGAATGTTCAATCTCTCAGAGCCCTGACCCCAATGTCTCTGAGACTCAGGGTCTGCTTCCC 180
 |||||
 1402 CACATTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAACAATTTCCAAAACGTGCA 1461
 |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:13, Search time 776.948 Seconds
(Without alignments)
10530.153 Million cell updates/sec

Title: US-09-936-271b-13_COPY_1_5000

Perfect score: 5000

Sequence: 1 gggccagagtgaggaagcag.....tgatccacagtgctgctgc 5000

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued Patents NA:
1: /cgn2_6/ptodate/1/ina/5A.COMB.seq:
2: /cgn2_6/ptodate/1/ina/5B.COMB.seq:
3: /cgn2_6/ptodate/1/ina/6A.COMB.seq:
4: /cgn2_6/ptodate/1/ina/6B.COMB.seq:
5: /cgn2_6/ptodate/1/ina/PTUS.COMB.seq:
6: /cgn2_6/ptodate/1/ina/backfile1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	44.9	10818	4 US-09-949-016-13583	Sequence 13583, A
2	444	8.9	246240	2 US-08-724-394A-20	Sequence 20, Appl
3	444	8.9	246240	2 US-08-724-394A-21	Sequence 21, Appl
4	444	8.9	246240	2 US-08-724-394A-22	Sequence 22, Appl
5	441	8.8	97376	4 US-09-949-016-16093	Sequence 16093, A
6	421.2	8.4	22973	4 US-09-949-016-13644	Sequence 13644, A
7	417.6	8.4	120727	4 US-09-949-016-15787	Sequence 15787, A
8	417.6	8.4	120727	4 US-09-949-016-15788	Sequence 15788, A
9	415.4	8.3	166698	4 US-09-949-016-16038	Sequence 16038, A
10	414.2	8.3	79858	4 US-09-949-016-16080	Sequence 16080, A
11	414	8.3	60304	4 US-09-949-016-12218	Sequence 12218, A
12	414	8.3	60305	4 US-09-949-016-15791	Sequence 15791, A
13	411.2	8.2	125188	4 US-09-949-016-11960	Sequence 11960, A
14	405.6	8.1	33379	4 US-09-949-016-13861	Sequence 13861, A
15	404.8	8.1	50217	4 US-09-949-016-16051	Sequence 16051, A
16	404	8.1	58593	4 US-09-949-016-12232	Sequence 12232, A
17	402.6	8.1	53442	4 US-09-949-016-11921	Sequence 11921, A
18	402.6	8.1	53453	4 US-09-949-016-13370	Sequence 13370, A
19	402.2	8.0	64518	4 US-09-949-016-17289	Sequence 17289, A
20	400.8	8.0	55114	4 US-09-949-016-16792	Sequence 16792, A
21	400.8	8.0	87617	4 US-09-949-016-16551	Sequence 16551, A
22	400.2	8.0	81701	4 US-09-949-016-14891	Sequence 14891, A
23	400.2	8.0	94855	4 US-09-949-016-12264	Sequence 12264, A
24	399.8	8.0	17189	4 US-09-949-016-17465	Sequence 17465, A
25	399.6	8.0	130563	4 US-09-949-016-12273	Sequence 12273, A
26	399.6	8.0	131379	4 US-09-949-016-16050	Sequence 16050, A
27	399.4	8.0	39920	4 US-09-949-016-16316	Sequence 16316, A

C 28	397.2	7.9	79350	4 US-09-949-016-12467	Sequence 12467, A
C 29	397.2	7.9	79351	4 US-09-949-016-16275	Sequence 16275, A
C 30	396.6	7.9	45469	4 US-09-949-016-13398	Sequence 13398, A
C 31	396.6	7.9	87734	4 US-09-949-016-17521	Sequence 17521, A
C 32	396.4	7.9	19451	4 US-09-949-016-13695	Sequence 13695, A
C 33	395.8	7.9	26104	4 US-09-949-016-14045	Sequence 14045, A
C 34	395.8	7.9	77626	4 US-09-949-016-12608	Sequence 12608, A
C 35	395.6	7.9	39920	4 US-09-949-016-16316	Sequence 16316, A
C 36	395	7.9	43255	4 US-09-949-016-11909	Sequence 11909, A
C 37	394.8	7.9	9226	4 US-09-949-016-12596	Sequence 12596, A
C 38	394.8	7.9	9226	4 US-09-949-016-15721	Sequence 15721, A
C 39	394.6	7.9	49487	4 US-09-949-016-15721	Sequence 15721, A
C 40	393	7.9	133613	4 US-09-949-016-15651	Sequence 15651, A
C 41	392.8	7.9	58782	4 US-09-949-016-11959	Sequence 11959, A
C 42	391.6	7.8	69834	4 US-09-949-016-11959	Sequence 11959, A
C 43	391.6	7.8	69834	4 US-09-949-016-12925	Sequence 12925, A
C 44	391.4	7.8	8405	4 US-09-949-016-11882	Sequence 11882, A
C 45	391.4	7.8	8405	4 US-09-949-016-15376	Sequence 15376, A

ALIGNMENTS

RESULT 1					
US-09-949-016-13583					
Sequence 13583, Application US/09949016					
Patent No. 6812339					
GENERAL INFORMATION:					
APPLICANT: VENTER, J. Craig et al.					
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED					
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF					
FILE REFERENCE: C0001307					
CURRENT APPLICATION NUMBER: US/09/949, 016					
CURRENT FILING DATE: 2000-04-14					
PRIOR APPLICATION NUMBER: 60/241,755					
PRIOR FILING DATE: 2000-10-20					
PRIOR APPLICATION NUMBER: 60/237,768					
PRIOR FILING DATE: 2000-10-03					
PRIOR APPLICATION NUMBER: 60/231,498					
PRIOR FILING DATE: 2000-09-08					
NUMBER OF SEQ ID NOS: 207012					
SOFTWARE: FastSeq for Windows Version 4.0					
SEQ ID NO 13583					
LENGTH: 10818					
TYPE: DNA					
ORGANISM: Human					
US-09-949-016-13583					
Query Match					
Best Local Similarity 44.9%; Score 2245; DB 4; Length 10818;					
Matches 2245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	2756	TGACGTGTCCTCTGTCAGGCGCATTATGATTTGGCTGAGTGCCTTATGATGC	2815		
DB	1	TGACGTGTCCTCTGTCAGGCGCATTATGATTTGGCTGAGTGCCTTATGATGC	60		
QY	2816	CCGATATTTGACCGGTGTGACTACCTGAACTCTGTGTAGGGTGAATGTGACTGT	2875		
DB	61	CCGATATTTGACCGGTGTGACTACCTGAACTCTGTGTAGGGTGAATGTGACTGT	120		
QY	2876	GTCGTCTGTGTAGGCGGTAAATGCTACTATGATGTGATGAGCTGTGTCT	2935		
DB	121	GTCGTCTGTGTAGGCGGTAAATGCTACTATGATGTGATGAGCTGTGTCT	180		
QY	2936	GGAGTTTCTGCTCTGCTGAGAGATAGAGGAGGTAGCTATCTTGGAGATG	2995		
DB	181	GGAGTTTCTGCTCTGCTGAGAGATAGAGGAGGTAGCTATCTTGGAGATG	240		
QY	2996	GGTCCAGGTGACTGACTGCTGAGTGTGCTGTGTGACAGAGATGTGCGACTGCA	3055		
DB	241	GGTCCAGGTGACTGACTGCTGAGTGTGCTGTGTGACAGAGATGTGCGACTGCA	300		
QY	3056	ACATCTGTGACACAGCGGATCTGTGTGTGACACTGACACTGTGTGATGAGGTGTGCG	3115		

[illegible]

Db	1381	TTTTTGTGTCTTAAAGTACAGAGGTTTACCTTGTGGCCAGGCTGTCTAACTCC	1440
Qy	4196	CAACCTAGGTGATCGGCCCACTCGGCTCCCAAGTGTGGAGTTACAGGCGTGAACC	4255
Db	1441	CAACCTCAGGTGATCGGCCCACTCGGCTCCCAAGTGTGGAGTTACAGGCGTGAACC	1500
Qy	4256	ACCGCCCCCAGGCCCAAGTCAAGACTTTTATATAGAGACTCTTAACATATTAACCTTGAACC	4315
Db	1501	ACCGCCCCCAGGCCCAAGTCAAGACTTTTATATAGAGACTCTTAACATATTAACCTTGAACC	1560
Qy	4316	TGGCCCTTAACATAGTCAATTCCAAACCCCTTCCGCTCCAGGCCCTGAACCCCACTCACTG	4375
Db	1561	TGGCCCTTAACATAGTCAATTCCAAACCCCTTCCGCTCCAGGCCCTGAACCCCACTCACTG	1620
Qy	4376	AGGCTTGACCCCACTTCTTGAAGACAGTTCCATCCCTAAAGCCCTGTCTCCCTCCATC	4435
Db	1621	AGGCTTGACCCCACTTCTTGAAGACAGTTCCATCCCTAAAGCCCTGTCTCCCTCCATC	1680
Qy	4436	CCCAAGGCTCCAGCCCCCAAGCTTTGGCACTAACCCCTGAGACTTGTCCAGGAATCCGTATC	4495
Db	1681	CCCAAGGCTCCAGCCCCCAAGCTTTGGCACTAACCCCTGAGACTTGTCCAGGAATCCGTATC	1740
Qy	4496	CCAATTTTAAACCCCTCAATGTAGTTCTTAAGCCAAATTCAGGAATCTGTAGAGGTCCAGTTAGA	4555
Db	1741	CCAATTTTAAACCCCTCAATGTAGTTCTTAAGCCAAATTCAGGAATCTGTAGAGGTCCAGTTAGA	1800
Qy	4556	GTCACGATACCCCTACCTGAGCCCTGGGCTCTGTCTCTTGAAGCTTGAGCTTGAGAGG	4615
Db	1801	GTCACGATACCCCTACCTGAGCCCTGGGCTCTGTCTCTTGAAGCTTGAGCTTGAGAGG	1860
Qy	4616	TGCCACTCTTAATTTCTCAAGGCCCTTGCCCCCTCAGACATGTACAGACACCCACTTTC	4675
Db	1861	TGCCACTCTTAATTTCTCAAGGCCCTTGCCCCCTCAGACATGTACAGACACCCACTTTC	1920
Qy	4676	TAGCTGGTCTGGGCTCTTGAAGTCTGAACACCCAGGCCCAAGCCCGGCTCTGAGCC	4735
Db	1921	TAGCTGGTCTGGGCTCTTGAAGTCTGAACACCCAGGCCCAAGCCCGGCTCTGAGCC	1980
Qy	4736	CCGCCCAACCCATTTTCCGTTCCAGAGACATGTTCTTGCCAAACATATGATGTTTCTGTGA	4795
Db	1981	CCGCCCAACCCATTTTCCGTTCCAGAGACATGTTCTTGCCAAACATATGATGTTTCTGTGA	2040
Qy	4796	CCACCCCTCTTAACACCGTGCCCTCTGGAGAACACAGAACTTGGAGCTTGGGGCCGGGGA	4855
Db	2041	CCACCCCTCTTAACACCGTGCCCTCTGGAGAACACAGAACTTGGAGCTTGGGGCCGGGGA	2100
Qy	4856	AGAGCGCCCGGTGGAGTGAACAGACAGCGGCATATCAATATGATCCGACTGGAGTATGCA	4915
Db	2101	AGAGCGCCCGGTGGAGTGAACAGACAGCGGCATATCAATATGATCCGACTGGAGTATGCA	2160
Qy	4916	CACCCACCGGTGGAGGACCGCGCTGTGCTTAAGAGCCCAACCAAGCTCTACTGCCGGCGGT	4975
Db	2161	CACCCACCGGTGGAGGACCGCGCTGTGCTTAAGAGCCCAACCAAGCTCTACTGCCGGCGGT	2220
Qy	4976	GTTGTGTGATTCACAGTGGCTGCTC	5000
Db	2221	GTTGTGTGATTCACAGTGGCTGCTC	2245
RESULT 2			
US-08-724-394A-20/c			
; Sequence 20, Application US/08724394A			
; Patent No. 5872237			
; GENERAL INFORMATION:			
; APPLICANT: Feder, John N.			
; APPLICANT: Krommal, Gregory S.			
; APPLICANT: Laufer, Peter M.			
; APPLICANT: Ruddy, David A.			
; APPLICANT: Thomas, Winston			
; APPLICANT: Tauchinashi, Zenta			
; APPLICANT: Wolff, Roger K.			
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1			
; TITLE OF INVENTION: Sequences and Antibodies Thereo			

```

1  NUMBER OF SEQUENCES: 31
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
4  STREET: Two Embarcadero Center, 8th Floor
5  CITY: San Francisco
6  STATE: CA
7  COUNTRY: USA
8  ZIP: 94111-3834
9
10 COMPUTER READABLE FORM:
11 MEDIUM TYPE: Floppy disk
12 OPERATER: IBM PC compatible
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14 SOFTWARE: Patentn Release #1.0, Version #1.30
15 CURRENT APPLICATION DATA:
16 APPLICATION NUMBER: US/08/724,394A
17 FILING DATE: 01-OCT-1996
18 CLASSIFICATION: 536
19 ATTORNEY/AGENT INFORMATION:
20 NAME: Plets, Renee A.
21 REGISTRATION NUMBER: 35,136
22 REFERENCE/DOCKET NUMBER: 017957-000100
23 TELECOMMUNICATION INFORMATION:
24 TELEPHONE: 415-576-0200
25 TELEFAX: 415-576-0300
26 INFORMATION FOR SEQ ID NO: 20:
27 SEQUENCE CHARACTERISTICS:
28 LENGTH: 246240 base pairs
29 TYPE: nucleic acid
30 STRANDEDNESS: not relevant
31 TOPOLOGY: not relevant
32 MOLECULE TYPE: CDNA
33 FEATURE:
34 NAME/KEY: misc feature
35 LOCATION: 1..246240
36 OTHER INFORMATION: /note= "HLA-H. CONTIG"
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967

```

Query Match 8.9%; Score 444; DB 2; Length 246240;

Best Local Similarity 80.1%; Pred. No. 5, 4e-93;
Matches 580; Conservative 0; Mismatches 120; Indels 24; Gaps 4;

Oy	335	TATGATCATCTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCGTCTCG	394
Dd	3310	TTTATTAAAGACTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAACGGAGTCTGTCTCG	3251
Oy	335	TCGCCAAGCTGAGTAGCAGTGGCGGAGATCTCGGCTCACTGCAGCTCCGCTCCCGGAT	454
Dd	3250	TCGCCAAGCTGAGTAGCAGTGGCGGAGATCTCGGCTCACTGCAGCTCCGCTCCCGGAT	3191
Oy	455	TCACGCCATTCTCTGCTCAAGCTCCCAAGTACTGGAGCTAACAGGGCGCCCGCACATAC	514
Dd	3190	TCACGCCATTCTCTGCTCAAGCTCCCAAGTACTGGAGCTAACAGGGCGCCCGCACATAC	3131
Oy	515	GCCCCGATAATTTTTTTGATTTTTTAAGTAGAGAAGGGGTTTCAACGTTTTAGCCGGAGT	574
Dd	3130	GCCCCGATAATTTTTTTGATTTTTTAAGTAGAGAAGGGGTTTCAACGTTTTAGCCGGAGT	3072
Oy	575	GCTTCGATCTCTGACTCTGTGATCCGCCCGCTCGGCTCCCAAGTCTGGAGTTACA	634
Dd	3071	GCTTCGATCTCTGACTCTGTGATCCGCCCGCTCGGCTCCCAAGTCTGGAGTTACA	3012
Oy	635	GGCGTAGACCAACCGGCGCCGAGCATGATCATCTTTCATCATAGTGTGATGTGACAAAGTAC	694
Dd	3011	GGCGTAGACCAACCGGCGCCGAGCATGATCATCTTTCATCATAGTGTGATGTGACAAAGT	2955
Oy	695	CTAAAGCATACGACTCTAACCTTTAAATATGCACTTTGGGCGACGACCGTGGCTCATG	754
Dd	2958	CTAATCTCTAATGTATG---TATTAAGATGAACCCCGGCTGGGCGAAGTGGCTACA	2903
Oy	755	CCTGTAATTCACGACTTTGGAGGACAGAGTGGGTGATCATCTTGAGGCCAGAGATTGG	814
Dd	2902	CCTGTAATTCACGACTTTGGAGGACAGAGTGGGTGATCATCTTGAGGCCAGAGATTGG	2843

QY	815	AGACGAGCTGGCCAAATGGTGAACCTGTCTTATATAAAAAAAAAAAAAAAAAA	874
Db	2842	AGACCAAGCTGGCCAAACGGCGAAACCTGTCTACT-----AAAATATCA	2795
QY	875	AAAATCAGCCGGGATGTGCTGGGGCAACCTGTAAATCCAGCTATGCTGGAGGCTGAGGCA	934
Db	2794	AAAATTGAGCTGGCATGTGTGGCAGGCAACTGTAAATCCAGCTGCTGTGGGAACCTGAGGCA	2735
QY	935	CGAAGATCACTTGAACCTGTGGAGGCGAAGGTGAGTGGGCGAGATCACTACCCGCC	994
Db	2734	GGAGAAATCGTTTGAACCTGTGGACGTGGGGATTTGAGTGAAGCCAAAGACTGCACACTGCAC	2675
QY	995	TTCAGCCTGGGCGACAGAGCAAGACTGTCTCAATTAATTAATTAACAAACGAAACAAG	1054
Db	2674	TTCAGCCTGGGCGACAGGTGAGACAACAAGTCTCAAAAAAATAAAAAATTAATAATTA	2615
QY	1055	CAGT 1058	
Db	2614	AAAT 2611	

RESULT 3

US-08-724-394A-21/c
; Sequence 21, Application US/08724394A

GENERAL INFORMATION:

1 APPLICANT: Feder, John N.
2 APPLICANT: Krommal, Gregory S.
3 APPLICANT: Lauert, Peter M.
4 APPLICANT: Ruddy, David A.
5 APPLICANT: Thomas, Winston
6 APPLICANT: Tsuchinashi, Zenta
7 APPLICANT: Wolff, Roger K.
8 TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e
9 TITLE OF INVENTION: Sequences and Antibodies Thereo
10 NUMBER OF SEQUENCES: 31

ADDRESSEE: TOWNSEND and CREW LLE

CITY: San Francisco

STATE: CA

COUNTRY: USA
7TP. 94111-3934

FILE: 94111-3634
COMPUTER READABLE

MEDIUM TYPE: F1

COMPUTER: IBM F

OPERATING SYSTEM: Patem

CURRENT APPLICATIONS

APPLICATION NUMBER

CLASSIFICATION: FILING DATE: 01

ATTORNEY/AGENT INFO

NAME: Fitts, Re

REGISTRATION NUMBER
REFERENCE/DOCKET

TELECOMMUNICATION.

TELEPHONE: 415-

TELEFAX: 415-57

; INFORMATION FOR SEQ
; SEQUENCE CHARACTER

LENGTH: 246240

TYPE: nucleic acid

TOPOLOGY: not known

MOLECULE TYPE: CD

FEATURE: NAME/REV. m100

NAME/ NO: 111BC
LOCATION: 1..24

OTHER INFORMATION:

US-08-724-394A-21

Query Match 8.9%; Score 444; DB 2; Length 246240;

QY 875 AAATGAGCGGGGTCGTGGGGGACACCTGTATCCAGCTATGCTGAGGCTGAGGCA 934
DB 2794 AAAATTGCTGGGCGATGTGACAGGACCTGTATCCAGCTGTTGGAGACTGAGGCA 2735
QY 935 CGAGAGTCACTTGAACCTGAGGCGGAGGTTGAGTGGGCGAGATCAATCAACCCGC 994
DB 2734 GGAAGATCGTTGAACCTGAGGCGGAGGTTGAGTGGGCGAGATCAATCAACCCGC 2675
QY 995 TCCAGCTGGGCGGACAGACAGACACTGTCTCAATTAATAAACAACAACAAG 1054
DB 2674 TCCAGCTGGGCGGACAGGAGGAGACAGCTCTCAAAAAATAAATAAATAAATAA 2615
QY 1055 CAGT 1058
DB 2614 AAT 2611

RESULT 5

US-09-949-016-16093
; Sequence 16093, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16093
; LENGTH: 97376
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(97376)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16093

Query Match 8.8%; Score 441; DB 4; Length 97376;
Best Local Similarity 76.5%; Pred. No. 1,7e-92;
Matches 584; Conservative 0; Mismatches 165; Indels 14; Gaps 3;

QY 347 CTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGCTGTGCGCCAGGCTG 406
DB 88600 CATTATTAATTTTTTTTTTTTTTTTTTTGAGTGTGAGTCTGTTGCCAGGCTG 88659
QY 407 GAGTGAGTGGCGGAGTCTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 466
DB 88660 AACACAGTGGCGGAGTCTGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 88719
QY 467 CCGGCTCAGGCTCCCAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 526
DB 88720 CCGGCTCAGGCTCTGAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG 88779
QY 527 TTTTGTATTTTGAAGAGACGGGGTTTACCGGTTTACCGGGATGGCTTGATCTCC 586
DB 88780 TTTTGTATTTTGAAGAGACGGGGTTTACCGGTTTACCGGGATGGCTTGATCTCC 88839
QY 587 TGAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 646
DB 88840 TGAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 88899
QY 647 CGGCGCGGCGGATGATCTCTTGTGACTGTGATGTGACAAATGAC-CTAAAGCATC 705

DB 88900 TGAACCGGCGCAATTTTAAATTAATGATCTTAAAGTCTCTCTGATCTTTT 88959
QY 706 AGACTCAACCTTTAAATATGAGTGTGGGCGGACCGTGGCTATGCTGTAATTC 765
DB 88960 TTTTGTATCTTAAACCTGTATTTTGGCGGACAGTACTCAATGCTCTTAATTC 89019
QY 766 AGCACTTTGGAGGAGGAGTGGGTGAATCACTTGAAGGCGAGGATTGAGACGAGCTG 825
DB 89020 AGCACTTTGGAGGAGGAGGCGGAGTGAATCACTTGAAGGCGAGGATTGAGACGAGCTG 89079
QY 826 GCCAAGTGTGAACCTGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 885
DB 89080 GCCAAGTGTGAACCTGTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 89127
QY 886 GGTGTGTGGGCGACACCTGTATCCAGTATGCTGAGGCTGAGGCGAGAGTACT 945
DB 89128 AGCGTGTGTGATGCTGTATCCAGTACTGAGGCGAGGCGAGAAATCTACT 89187
QY 946 TGAACCTGTGAGGCGGAGGTTGAGTGGGCGGAGATCAATCAACGCTCGAGCTGGG 1005
DB 89188 TGAACCGAGAGGCGAGGTTGAGTGGGCGGAGATGCTTCACTTGAATCTCGAGCTGGG 89247
QY 1006 CGAC-AGAGCAGACTGTCTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1064
DB 89248 TGAAGAGGAGAACTCAATCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 89307
QY 1065 TACCTATGTTATTTCTAAAAAATGCTGTCAACAATAGA 1107
DB 89308 TAGAGCAGTTATGTTTCAACAGCAAGTTGTACAGAAATGACA 89350

RESULT 6

US-09-949-016-13644
; Sequence 13644, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13644
; LENGTH: 22973
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13644

Query Match 8.4%; Score 421.2; DB 4; Length 22973;
Best Local Similarity 78.9%; Pred. No. 3.7e-88;
Matches 549; Conservative 0; Mismatches 123; Indels 24; Gaps 3;

QY 357 TTTTGTATTTTGAAGAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416
DB 17364 TTTTGTATTTTGAAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 17423
QY 417 GCGGATCTCGGCTCACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476
DB 17424 GCAATCTCGGCTCACTGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 17483
QY 477 CTTCCCAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 536
DB 17484 CTTCCCAAGTGTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 17543
QY 537 TTTAGTAGAGACGGGGTTTACCGTTTATGCGGAGTGGCTGAGTCTCTGACCTGCTG 596

```

Db      17544 TTTATGATAGAGACAGGGTTTCAACCGTGTAGCCAGAGATGGTCTTGATTTCTTGACCTTG 17603
QY      597 ATCCGCCCGCCCTCGGCTTCCCAAAGTCTTGGAATACAGGCTGAGCCACCGGCCGCGC 656
Db      17604 ATCCACTGCTCAGCTCCCAAAGTCTGGAGTTACAGCGTGAAGCCACGAGCGTGAC 17663
QY      657 CA----TGATCATCTTCTTGACTAGTCTGATGTA-----CAAGTACTTAAGCCA 703
Db      17664 CAACCTTATTTTCAGTCTTGCGCTGAGCTTTTATTACTTACTATGAAATTTGTAACATA 17723
QY      704 TCAGACTCAACCTTTAAATATGCAATTTGGGCGAGGACCGGAGCTCATGCTGTAATT 763
Db      17724 CACAAAAGTAAATAGAAATAGATATGATGAGGCGAGGTGCAAGTGCACACCTGTATC 17783
QY      764 CCAGACCTTTGGGAGGCGAGGCTGAGGTGATCATCTTGAGGCGAGATTTGAGACCGC 823
Db      17784 CTAGCACTTTGGGAGGCTGAGGTGATCTGATGAGCTTGAGATTTCTAATCCAGCC 17843
QY      824 TGGCCAACTGCTGTAATCTCTGTCTTTACTTAAAAAATTTAAAAAATTTAAATCAGC 883
Db      17844 TGGCCAACTGCTGTAATCTCTGTCTTTACTTAAAAAATTTAAAAAATTTAAATCAGC 17892
QY      884 CGGGTGTGTTGGGGGACACCTGTAATCCAGCTATGCTGAGGCTGAGGCGAGAGTCA 943
Db      17893 TGGATGTGTGCGCACACCTGTAATCTGAGCTACTCGGAGGCTGAGGCGAGAGTCA 17952
QY      944 CTTGAAACCTTGAGGGGAGGTTGAGTGGGCGGAGATCATCAACCGCCTCAGCCTG 1003
Db      17953 CTTGAACTCTGGAGGTTGAGGTTGAGGTGAGCCGAAATTCAGCCATTTGATTCTAGCCTG 18012
QY      1004 GGCGACAGAGCAAGACTCTGTCTCAATATAATTAAT 1039
Db      18013 GGCAACAGAGCGAGACTCAGTCTCAAAAAAAACT 18048

RESULT 7
US-09-949-016-15787/c
; Sequence 15787, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH THE INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15787
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15787

```

Query Match	8.4%	Score 417.6	DB 4	Length 120727
Best Local Similarity	80.5%	Pred. No. 5.8e-87		
Matches 546	Conservative 0	Mismatches 109	Indels 23	Gaps 4
OY	362	TTTTTTTTTTTTTTGAACGAGTCTGCTCTGTGCGCCAGCGCTGAGATGCACTGGCGGG	421	
Db	32483	TTTTTTTTTTTTTTGGACACGAGAGTCTGCTCTGTGACCCGAGCGTGAAGATGCAATGCTGGG	32424	

OY	422	TTCGCGGTCACTGCAGATCCGGCTCCCGGGTTACAGCATTCCTCCTCAGCTCC	481
Db	32423	ATTCCGGTCACTGCAGAGTCCGGCTCCCGGGTTACAGCATTCCTCCTCAGCTCC	32364
OY	482	CAAGTAGCTGGACTACAGCGGCCGCCACTAOCGCCGCTAA-TTTTTTGTAATTTTA	540
Db	32363	CAAGTAAACAAGGAATAAGAGGGCCCCGACCAACGCCCTGGCTAATTTTTTTTGTAATTTTA	32304
OY	541	GTAAGACAGGGGTTTACCGTTTTAGCCGGGAATGGCTCTGATCTCCTGACCTCTGATCC	600
Db	32303	GTAGAGAAGGGGTTTCACTGTGTGTCAGAGATGTGTTGATCTCTCTGACCTCTGATCC	32244
OY	601	GCCCCCTCCGGCTCCCAAAGTGTGGGATTTACAGGCTGAGCACACGGCCGGGCATG	660
Db	32243	ACCACCTCGGCTCTCCAAAGTGTGGGATTTACGGGCTGAGCACACCGCCCGGCCAGC	32184
OY	661	ATCATCTTCTTGACTATGCTGATGTGACAAAGTAAAGCATCAGACTTACCTTTTA	720
Db	32183	CATGTCGAGTTCTGT- TGATGTTCTTCACTCTGAAGGCCACGTGAAGAGGCTT--	32127
OY	721	AATATGACGTTTGGGCCAGGACCCGTGGCTCATGCTGTAAATTCCAGGACTTTGGGAGGC	780
Db	32126	-----AGACTTGGCTGTCACGGTGGCTCAACCTGTTCATCTAGCACTTTGGAGAGC	32074
OY	781	AGAGGTGGGTGAATCACTTTGAGGCGCAGAGTTTGAGACCAAGCTGGCCAACATGTGAAA	840
Db	32073	CGAGGCAAGTAGATCACTGTAGGCCAGAGAGTTGAGAACAAAGTGTGACCAACATGTGAAA	32014
OY	841	CTGTGCTTTTACTAAAAAATAAAAAAAAAATCAGCCGGGTGTCTGTGGGAC	900
Db	32013	CCCCCTCTCTACT-----AAAACTCAAAAAATTAAGTGGGCTGTGGGTGGCAC	31966
OY	901	ACCTGTATCCAGCTATGCTGAGAGCTGAGGCGCAGAGGTCACTTGAACCCCTGAGAGCG	960
Db	31965	ACCTGTATCCAGCTTACTCAGAGAGCTGAGGCGCAGAGGATCTTTGAATCTCAGAGGCA	31906
OY	961	GAGGTTCAGTGGGCCGAGATCACTCACCGGCTTCCAGCCTGGGCGACAGACAAAGACT	1020
Db	31905	GAGGTTCAGTGAAGCCAAAGATTGGCCACTGCATCTCAAGCCTGGGCGACAGACCAAGACT	31846
OY	1021	CTGTCTCAAAATTAATTAATTA 1038	
Db	31845	CTGTCTCAAAAAAATAA 31828	

```

RESULT 8
US-09-949-016-15788/C
; Sequence 15788, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15788
; LENGTH: 120727
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ..(120727)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15788

```


Query Match 8.4%; Score 417.6; DB 4; Length 120727;

Best Local Similarity 80.5%; Pred. No. 5.8e-87;

Matches 546; Conservative 0; Mismatches 109; Indels 23; Gaps 4;

```

QY 362 TTTTGTGTTTGTGAGCAGAGTCTGCTCTGTCGCGCCAGAGCTGAGTGAAGGCGG 421
DB 32483 TTTTGTGTTTGTGAGCAGAGTCTGCTCTGTCGCGCCAGAGCTGAGTGAAGGCGG 32424
QY 422 ATCTCGGCTCAGTCAAGTCTCCGCTCCGCGTTTACGCCATTCTCTGCTCAGCTCC 481
DB 32423 ATCTCGGCTCAGTCAAGTCTCCGCTCCGCGTTTACGCCATTCTCTGCTCAGCTCC 32364
QY 482 CAAGTACGTGAGTCAAGCAGCGCCGCTCAGTCAAGCAGCGCGCTTAA-TTTTGTGTTT 540
DB 32363 CAAGTACGTGAGTCAAGCAGCGCGCCGCTCAGTCAAGCAGCGCGCTTAA-TTTTGTGTTT 32304
QY 541 GTAGAGAGCGGGGTTTACCGTTTATGCGGGATGGCTCGATCTCTGACCTCGTATCC 600
DB 32303 GTAGAGAGCGGGGTTTACCGTTTATGCGGGATGGCTCGATCTCTGACCTCGTATCC 32244
QY 601 GCTCGCTCGGCTCCGCAAGTGTGGATTACAGCGGTGAGCCAGCGCGCGCGCATG 660
DB 32243 ACCCAGCTCGGCTCCGCAAGTGTGGATTACAGCGGTGAGCCAGCGCGCGCGCATG 32184
QY 661 ATCATCTTCTGACTAGTCTGATGTGCAAGTACTTAAAGCATCACTCTACCTTTA 720
DB 32183 CATGTCTGATCTGT-TGATGTCTTCTTCACTTAAAGCATCACTTAAAGCATCTT 32127
QY 721 AATATGAGTTTGGGCGCAGGACCGTGGCTCATGCTGTATTTACAGACTTTTGGAGGC 780
DB 32126 -AGACTTGGCTGTGCTGAGTGTGCTCACTGTCTATCTTACACTTTTGGAGGC 32074
QY 781 AGAGTGTGTGAATCACTTGAAGCGCAGAGTTTGAACAAGCTGTGCAATGTGAAA 840
DB 32073 CGAGGCGGTGATGATCACTTGAAGCGCAGAGTTTGAACAAGCTGTGCAATGTGAAA 32014
QY 841 CTCTGTCTTAACTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 900
DB 32013 CCGGCTCTTAACTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 31966
QY 901 ACCGTGATCCGAGCTATGCTGAGGCTGAGGCGCAGAGTCACTTGAACCTTGGAGGC 960
DB 31965 ACCGTGATCCGAGCTATGCTGAGGCTGAGGCGCAGAGTCACTTGAACCTTGGAGGC 31906
QY 961 GAGGTGACGTGGCGCAGATCAATCAACCGCTTCCAGCTTGGCGCAGAGCAAGACT 1020
DB 31905 GAGGTGACGTGGCGCAGATCAATCAACCGCTTCCAGCTTGGCGCAGAGCAAGACT 31846
QY 1021 CTGTCTCAATAATAATAA 1038
DB 31845 CTGTCTCAATAATAATAA 31828

```

RESULT 9

```

US-09-949-016-16038
; Sequence 16038, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0

```

SEQ ID NO 16038

LENGTH: 16698

TYPE: DNA

ORGANISM: Human

FEATURE:

NAME/KEY: misc feature

LOCATION: (1)-(16698)

OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16038

Query Match

Best Local Similarity 76.2%; Pred. No. 2.2e-86;

Matches 560; Conservative 0; Mismatches 151; Indels 24; Gaps 3;

```

QY 328 TGTGTCTATGATCTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 387
DB 106710 TACCTGTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 106769
QY 388 CGCTGTGTGCGCAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 447
DB 106770 TGTCTGTGTGCGCAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 106829
QY 448 CCGGGTTTCAAGCATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 507
DB 106830 CTGTGTTTCAAGCATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 106889
QY 508 CCACTACGCGGCTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 567
DB 106890 CCACTACGCGGCTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 106949
QY 568 CGGATGTGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 627
DB 106950 CAGATGTGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 107009
QY 628 GATTACAGCGGTGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 687
DB 107010 GATTACAGCGGTGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 107069
QY 688 CAAGTACTTAAAGC-----ATCAGACTTACCTTTAAATATGACGTTGGGC 736
DB 107070 AAAAAAATATGCAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 107129
QY 737 CAGGCA-CGCTGTCTATGCTGTATTTTCAAGACTTTTGGAGGCGAGAGTGTGTAATC 795
DB 107130 TTTAAAGTGTGCTCAAGCTGTATTTTCAAGACTTTTGAAGGCTGTAACCGGCGAGATC 107189
QY 796 ACTTGAGGCGCAGAGTTTGAACAAGCTGTGCAATGATGTTGTTGTTGTTGTTGTTGTTGTT 855
DB 107190 ATTGAGGCTGAGAGTTTGAACAAGCTGTGCAATGATGTTGTTGTTGTTGTTGTTGTTGTT 107247
QY 856 AAAAAAAAAAAAAAAAAAATCAAGCGGCTGTGTGTGGGCGACACTGTAAATCCAGC 915
DB 107248 -----AAAAATACAAAATTTAGCCAGCGGTGTGTCACACGCTGTAAATCCAGC 107297
QY 916 TATCTGAGGCTGAGGCGCAGAGTCACTTGAACCTTGAAGCGGAGGTTGAGTGGCG 975
DB 107298 AACTCAGAGGCTGAGGCGCAGAGTCACTTGAACCTTGAAGCGGAGGTTGAGTGGCG 107357
QY 976 CGAGATCAATCAACCGCTTCAAGCTTGGCGCAGAGCAAGCTGTCTCAATAAT 1035
DB 107358 TGAGATCAATCAACCGCTTCAAGCTTGGCGCAGAGCAAGCTGTCTCAATAAT 107417
QY 1036 AATTAACCAACGAA 1050
DB 107418 AAAAAAAAAAAAAA 107432

```

RESULT 10

```

US-09-949-016-16080
; Sequence 16080, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

```

```

?
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
?
? FILE REFERENCE: CLO01107
?
? CURRENT APPLICATION NUMBER: US/09/949,016
?
? PRIOR FILING DATE: 2000-04-14
?
? PRIOR APPLICATION NUMBER: 60/241,755
?
? PRIOR FILING DATE: 2000-10-20
?
? PRIOR APPLICATION NUMBER: 60/237,768
?
? PRIOR FILING DATE: 2000-10-03
?
? PRIOR APPLICATION NUMBER: 60/231,498
?
? PRIOR FILING DATE: 2000-09-08
?
? NUMBER OF SEQ ID NOS: 207012
?
? SOFTWARE: FastSeq for Windows Version 4.0
?
? SEQ ID NO: 15080
?
? LENGTH: 79858
?
? TYPE: DNA
?
? ORGANISM: Human
?
? US-09-949-016-16080

```

Query Match	8.3%	Score 414.2	DB 4	Length 79858
Best Local Similarity	76.2%	Pred. NO. 3e-86		
Matches 557, Conservative	0	Mismatches 128	Indels 46	Gaps 2

OY	349	TTTTTTTTTTTTTTTTTTTTTTTTTTTGAACGGAAGTCGCTCTGTGCGCCAGGCTGGA	408
Db	10498	TTTTTTTTTTTTTTTTTTTTTTTTTTTGAACGGAAGTCGCTCTGTGCGCCAGGCTGGA	10557
OY	409	GTGCAAGTGGCGGAGATCTCGGCTCACTGCAGAGCTCCGGCTTCCAGCCATTCTCC	468
Db	10558	GTGCAATCGGCGTATCTTGGCTCACTGCAGAGCTCAATCTCCGGGTTCAATCATTTCTCC	10617
OY	469	TGCTCAAGCTCTCCCAAGTAGTGGAGACTACAGCGCCCGCATTAAGCCCGCTAATTTT	528
Db	10618	TGCTCAAGCTCTCCCAAGTAGTGGAGACTACAGCGCCCGCATTAAGCTTAAATTTT	10677
OY	529	TTTGTATTTTTTAGTAGAGACGGGGTTTCAACGTTTTTAGCCGGAGTAGGCTCAATCTCTG	588
Db	10678	TTTGTATTTTTTAGTAGAGACAGGGTTTCAACATGTTAGCAGGCTGTCTCAATCTCTCG	10737
OY	589	ACCTGTGATCCGCCGCTCGGCTCCCAAAGTGTGGGATTAACAGCGGTGAGCCACCG	648
Db	10738	ACCTGTGATCCGCCGCTCTGCTCTCCAAAGTGTGGGATTAACAGCGGTGAGCCACCG	10797
OY	649	CGCCCGGCAGTATCATCTTCTTGACATATCTGATGTGCAAGT-----	692
Db	10798	CGCCCGGCAGGAATCTTCTTCTTCACTTAAACAAGATCAATAGTCATCTATATTTTC	10857
OY	693	-----ACCTAAGCATCAGACTCTACCTTTAATAATGACAGTTTGG	734
Db	10858	TTCTAATAAAACCTGATGATGATATATATATCATACATAGATATTTAATACATCAAAAGG	10917
OY	735	GCCAGGACCGTGGCTCATGCTGTAAATTCCAGCACTTTGGGAGCAGAGTGGGTGAAT	794
Db	10918	GCCAGATGGCGTGGCTCATGCTGTAAATCCAGCACTTTGGGAGCGTGGCGGGTGAAT	10977
OY	795	CACTTGAGGCGCAGAGTTTAGACACAGCTCGGCGCAACATGGTGAATCTGTCTTACTTA	854
Db	10978	CACCTGAGGTCAAGATTCAAGACAGCTCGGCGCAAGATGTATAATCCCGTCTTACTT-	11036
OY	855	AAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTGTGGGCGACACCTGTAAATCCAG	914
Db	11037	-----AAAAATACAAAAAATTTAGCAGGCGTGTGTGGCGACCTGTAAATCTCA	11085
OY	915	CTATGCTGAGGCTGAGAGCGACGAGAGTCACTTAAACCTGTGAGAGCGGAGTTGCAATGGG	974
Db	11086	CTACTGGGGTGACTGAGAGCGAGGAATCGCTTAAACCCAGAGAGCGAAGATTTCCGTGAG	11145
OY	975	CCGAGATCAATCACTCCCTCTCAGACTGTGGGCAACAGAGCAAGTCTGTCTCAATTA	1034
Db	11146	CCAGATCAACGCCACTGCACTCTCGCTGGGCGATGAGTGAAGGCTCTGTCTCAAAAAA	11205
OY	1035	TAAATTAACA	1045

```
DB      11206 AAAAAATAAAA 11216

RESULT 11
US-09-949-016-12218/c
; Sequence 12218, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/231,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12218
; LENGTH: 60304
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12218

Query Match          8.3%; Score 414; DB 4; Length 60304;
Best Local Similarity 78.5%; Pred. No. 2.9e-86;
Matches 551; Conservative 0; Mismatches 135; Indels 16; Gaps 4;

QY      352 TTTTYYYYTTTTTTTTTTTTTTTTTTGAGACGAGCTCGCTGTGCGCCAGGCTGGAGTG 411
DB      36261 TTGTCCTTTTTTTTTTTTTTTTTTTTGAATGAGAGT-----CTCATGTGCTAGGCTGGAGTT 36206

QY      412 CAGTGGCGGAGATCCTCCGCTCACTGAACTCCGCTCCCGGGTTACAAGCATTCCTCTG 471
DB      36205 CAGTGGCGGAAATCTCGGCTCACTGAAAGTCCGCTCCCGGGTTACAGCATTCCTCTG 36146

QY      472 CTCAGCCTCCCAAAGTAGCTGGAGACTACAGGCGCGCCACTACGCCCCGCTAATTTTTT 531
DB      36145 CTCAACCTCCAGAAGTAGCTGGAGACTACAGGCGCCCAACCAACGCCCCGCTAATTTTTT 36086

QY      532 GTAATTTTAGTAGAGACGGGGTTTACAAGCTTTTACCGGGATGGCTCGATCTCTGACC 591
DB      36085 GTATTTTATAGTAGAGACAGGGGTTTACCGGTGTACCAAGATGGTCTCATATCTCCAGACC 36026

QY      592 TCAGTATCCGCGCCGCTCGGCTCCCAAGTCTGGGAATTACAGGCGTAGCACCGCGC 651
DB      36025 TCATBACCCGCTGCTCATGCTCCCAAGTCTGGGAATTACAGGCGTAGAGTACG-CAC 35967

QY      652 CCGGCGATGATCATCTTCTTGAATATGCTGATGTGACAAATCACTTAAGCCATCAGACTT 711
DB      35966 CGTGCCTGGCCCTGCCCAAGGCTGGTTTCTATGCTCTCAAGCCTTCATCATCATCTGG 35907

QY      712 TAACCTTAATATATGACATTTTGGGCGAGGACCGTGGCTCATAGGCTGTAATCCAGACT 771
DB      35906 AACCTTGCAAGAGCTGATTAAGGCGAGGACAGTGGCTCATATCAATATATCCAGACT 35847

QY      772 TTGGAGAGCAGAGTGGGAGTATCACTTGAGGCGAGATTTGAGACCAAGCTGGCCAAC 831
DB      35846 TTGGGATCTGATGAGTGGATGATCACTTGAATGACAGAGTTCAAGACCGCTGACCAAC 35787

QY      832 ATGTGAATCTGTCTTACTTAATAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTTC 891
DB      35786 ATGTGAATCTCCATCTCTACT-----AAAAATCAAAAAAATTAAGCAGCGCGTA 35737

QY      892 GTGGGGGACACCTGTAATCCAGCTATGCTGGAGGCTGAGGACGAGAGTCACTTGAACC 951
DB      35736 GTGGTACACGCTGTATATCCAGCTACTTGAAGAGGTGAGGACGAGGAATCACTTGAACC 35677

QY      952 CTGAGAGCGGAGGTTGACAGTGGGCGGAATCAATCACCGCCCTCCAGGCTGGGCGACAG 1011
```

Query Match	8.3%;	Score 414;	DB 4;	Length 60304;
Best Local Similarity	78.5%;	Prod. No. 2, 9e-86;		
Matches 551;	Conservative	0;	Mismatches 135;	Indels 16;
			Gaps	4;
OY	352	TTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGAGCTTCGCTGTCGTGCCACAGGCTGAGAG	411	
Db	36261	TTTTCTTTTTTTTTTTTTTTTTTTTTTGGATGAGAGT-----CTGTGCTGAGGCTGAGT	36200	
OY	412	CAGTGGGGGAGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTACAGCATCTCTGTC	472	
Db	36205	CAGTGGGGCAATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTACAGCATCTCTGTC	36144	
OY	472	CTCAGCCTCCCAAGTAGCTGGAGCTACAGGCGCCGCGCACTAAGCCCGCTAAATTTTTTT	531	
Db	36145	CTCAACCTCCAGAGTAGCTGGAGCTAAGCGCCGCGCAACACAGCCGGCTAAATTTTTTTT	36088	
OY	532	GTATTTTTTAGTAGAGACGGGGTTTACACCGTTTTTATGCGGGATGGCTCTCATCTTCGACC	591	
Db	36085	GTATTTTTTAGTAGAGACGGGGTTTACACCGTTTTTATGCGGGATGGCTCTCATCTTCGACC	36022	
OY	592	TGCGATCCGCGCCGCTCCGCGCTCCCAAGTGTCTGGGATTACAGGCGTGAAGCAACGCGC	651	
Db	36025	TGCGATCCGCGCTCCGCGCTCCCAAGTGTCTGGGATTACAGGCGTGAAGCAACGCGC	3596	
OY	652	CCGGCCATGATCATCTTCTTGAATGCTGATGTGACAGTACCTTAAAGCATGACATC	711	
Db	35966	CGTCCCTGCGCCGCGCCCAAGGCTGTTTTATGCTCTCAAGCCTTCGTATGCCACTCTGG	35907	
OY	712	TACCTTTAAATATGACGTTTGGGCGAGGACCGTGGCTATGCTGTATTTCCAGACT	772	
Db	35906	AACCTTTTCAAGACTGATTGAGCCAGGACAGTGGCTATATCTTAAATCCAGACT	3584	
OY	772	TTGGAGGAGAGAGTGGGTGAATCATCTTGAAGCCAGAGATTGAGACCAAGCTGGCCAC	831	
Db	35846	TTGGAGAGCTGAGGTGATGGATCATCTTGAAGTTCAGAGTTCAAGATCAAGCTGACCAAC	3578	
OY	832	ATGCTGAACCTCTGTCTTTTACTTAAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTCTC	891	
Db	35786	ATGCTGAACCTCCATCTCTACT-----AAAAATACAAAAAAATTTGACAGGGGTA	3573	
OY	892	GTGGGGGACACCTGTATCCAGCTATGCTGGAGGCTGAAGGACGAAGTCACTTGAAC	951	
Db	35736	GTGGTACACGCTGTATATCCAGCTATTGAGAGGCTGAAGGACGAAGATCATTTGAAC	3567	
OY	952	CTGAGGCGGAGGTTCAGTGGGCGAGATCATATCACCGCCCTCCAGCTTGGGCGACAG	1011	

us-09-936-271b-13 copy 1 5000.rnd

QY	892	GTGGGGGACAC	CTGTATATCCAC	ACTATGCTGAGG	GTGAGGACACG	AGATCACTTGA	AACC	951
QY	35736	GTGGTAAACAC	CCCTGTATATCC	ACACTCTTAGG	AGGCGCTGAGG	ACAGGAATCACTT	GAAC	35677
QY	952	CTGGAGCGGAG	GTTGAGTGGG	CGAGATCATAC	ATACCGCCCTC	CAGGCTGGG	CGACAG	1011

```

DB
35617 AGAGGACTGTCTCAAAAACAAAACAAAACAAAACCA 35576

RESULT 13
US-09-949-016-11980/c
; Sequence 11980, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0

```

```
;
; TYPE: DNA
; ORGANISM: Human
```

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) _ (125188)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11980

Query Match      8.2%; Score 411.2; DB 4; Length 125188;
Best Local Similarity 78.7%; Pred. No. 1,9e-85;
Matches 535; Conservative 0; Mismatches 128; Indels 17; Gaps 3;

Oy      347  CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACGAGCTCGCTCTGCGCCACGCTG 406
          |||||
Db       120787 CTTTTTTTTCTTTCCTTTTTTTTTTTTTTTTTTTTGGAGAAAGAGCTTGCTTTTGCCACGAGCCA 1207288

Oy      407  GAGTGCATGCGGGATCTCGGCTCACGCAAGCTCCGCTCCCGGGTTACGGCCATTCT 466
          |||||
Db       120727 GACTGCATGCGACTATCTCGGCTCACGCAAGCTCCGCTCTGGGTTACGCCATTCT 1206688

Oy      467  CCTGCTCAGGCTCCCAAGTAGCTGGAGCTACAGGCGCCGCGACCTACGCCCGGCTAAATT 526
          |||||
Db       120667 CATGCTCAGGCTCCCAAGTAGCTGGAGTTACAGGTGCTGCGCACCGTGCGGCTAA-T 1206099

Oy      527  TTTTTTGTATTTTGTAGTAGAGACGGGGTTTCAACGTTTTTAGCCGGAGTGGCTCGATCTCC 586
          |||||
Db       120608 TTTTGTATTTTGTAGTAGAGACGGGGTTTCAACGTTTAAACMAATGTGCTCGATCTCC 1205499

Oy      587  TGACCTCGTGATTCGGCGCGGCTCTGGGCTCCCAAAATGCTGGGATTACAGGCGGTGAGCCAC 646
          |||||
Db       120548 TGACCTCGTGATTCACCGCGCTCTGGGCTCCCAAAATGTCTAGGATTACAGGCGGTGAGCCAC 1204899

Oy      647  CGCGCCCGGCGCATGATCATCTTCTTGACTATGCTGATGTGACAAATACCTAAAGCATCA 706
          |||||
Db       120488 CGCGCCCGGCGCATGCGAGACTTCTTGAGTATCATATGTGGCTCTCTGTAAATTAATCT 1204299

Oy      707  GACTTCACCTTTTAAATATGACGTTTGGGCCACAGGACCGTGGCTCATGCTGTAAATTCCA 766
          |||||

```

QY	767	GCACCTTTGGGAGGACGAGGTTGGGTAACTCACTTTGAGCCAGGACTTTGAGACGAGCTGG	826
Db	120372	GCACCTTTGGGAGGCCAGGCGGACAGATACACTTGAAGTTGAGAGTTGAGACCAAGCTGG	120313
QY	827	CCAACATGCTGTAACCTCTGTCTTAACTAAAAAAAAAAAAAAAAAAAAAAATCAGCCGG	886
Db	120312	CCAACATGCTGAACCCCTATTCTACT-----AAAAATTCAAAAAATACACTGG	120265
QY	887	GGTGTGTGGGGGACACCTGTAATCCAGCTATGCTGGAGGCTGAGGCGACGAGAGTCACTT	946
Db	120264	GGGGGGTGGCAGCGCTGTATGTCGCCAGCTATTGGGAAGCTTAACTGAAGATACATT	120205
QY	947	GAACCCCTGGAGGCGGAGGTTGCAGTGGGCGCGAGATCACATCAACCGCCCTTCAGCCTGGGC	1006
Db	120204	GAACCCCGGAGGTTGGAGTTGCAGTGAGCCAAAGATGCGACGACTGTAAACCCGAGCCTGGGGT	120145
QY	1007	GACAGAGCAAGACTCTGTCT	1026
Db	120144	GACAGAGCGAGACTCTGTCT	120125

```

RESULT 14
US-09-949-016-13861
; Sequence 13861, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13861
; LENGTH: 33379
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(33379)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13861

```

Query Match	8.1%	Score 405.6	DB 4	Length 33379
Best Local Similarity	70.0%	Pred. No. 2e-84		
Matches 633	Conservative	0	Mismatches 209	Indels 62
			Gaps 4	
QY	335	TATGATCATCTCTCTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGGAGTCTCGCTCTG	394	
Db	5215	TATGGTGTAAATTAATTAATCTTTTTTTTTTTTTTTTTTTTTTTTGTGACGAGTCTCGCTCTG	5274	
QY	395	TCGCCACGACTGAGTGCAGTGGCGGAACTTCGAGCTCACTGCMACTCCGCTCCCGGCT	454	
Db	5275	TCGCCACGACTGAGTGCAGTGGCGGAACTTCGAGCTCACTGCMACTCCGCTCCCGGCT	5334	
QY	455	TCAGGCCAATTCCTCCGCTCAGCCTCCCAAGTGTGGACATACAGGCGCCCGCATACAC	514	
Db	5335	TCAGGCCAATTCCTCCGCTCAGCCTCCCAAGTGTGGACATACAGGCGCCCGCATACAC	5394	
QY	515	GCCCGCATTAATTTTGTGATTTTGTAGTAGAGACGGGCTTTCACCGTTTAGCCGGAGATG	574	
Db	5395	GCCCGCATTAATTTTGTGATTTTGTAGTAGAGACGGGCTTTCACCGTTTAGCCGGAGATG	5455	
QY	575	GCTTCGATCTCTGACCTGTGTATCCGCCGCTCGGCTCCCAAAAGTGTGGATTACA	634	

Db	5454	GTCTCGAATCTCTGACCTCGAGTCGCGCGCGCTCAGCGCTCCAAAGNCGTGGGATTACA	5513
OY	635	GGCGTAGGCCACCGCGCGCGCCGACATGATCATCTTCTT	671
Db	5514	GGTGTGAGCCACCGCGCGCGCGCGGTGTAAATTTATCTTAGCCATGTGTTCAAGTGA	5573
OY	672	-----GACTATGCTGATGTGACAGTACCTTAAAGCCATCA	706
Db	5574	ACAAAGTTGTACATGAAGACGTTTAGTAAATTTTGTAAATTTAAATATGTATGTA	5633
OY	707	GACTCTACCCCTTTAAATATGCAAGTTTG36GCGACCGTGGCTCATGCTGTAAATTC	766
Db	5634	GATATGTAGTGTTTTAAAGAGTTTCAGGGCGCGCGGTGGCTCAGCGCTGATCCCA	5693
OY	767	GCACCTTTGGAGGCGAGAGTGGGTGTGATTCATCTTAGGCGCAGAGTTTGAAGCCAGCTCG	826
Db	5694	GCACCTTTGGAGGCGCGAGGCGAGGTGATCAC--GAGGTCAAGATATCAAGCCATCTCG	5751
OY	827	CCAACATGATGAACTCTGTCTTACTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTT	886
Db	5752	CTAACATGATGAAATCCCATCTCACT-----AAAAATGCAAAAAATTAAGCCGG	5800
OY	887	GTGTCTGTGGGCGACACTGTATATCCAGCTATGTCTGAGGCTGAGGCGCAGAGTCACTT	946
Db	5801	GTGTGTGTGGCGACGCGCCTTATATCCAGCTACTGTGAGGTGAGGCGAGTAATGCGAT	5860
OY	947	GAACCCGTGGAGCGGAGGTGTGCACTGGGCGAGATCAATCAACGCGCCTCAGGCTG36G	1006
Db	5861	GAACCTTGGAGGCGGAGACTGTGAGTGAAGCGAGATCTGTGCACTGCACTCCAGCTG66G	5920
OY	1007	GACAGAGCAAGACTCTGTCTCAATTAATAATAATAACAAAGAAAGAAAGAGTTGTTGTA	1066
Db	5921	GACAGAGCGAGACTGTGTCTCAAAAAAATTTTTTTTTTTCAATAAATG36GCT	5980
OY	1067	CCTTAGTTATATCTAAAAAAAATGCTGTCAACAATPAAGCGAGAGTGAATTAAGA	1126
Db	5981	CAAAAGTAGACCTTTTGATGTATTAATGTTTATTTTCTTTAATGCGAGCTTCATAG	6040
OY	1127	AAATTAATGGCGCAAGAACTCTAAGGTATATTTTGACAAATCATTCAGAACTTTAAAAA	1186
Db	6041	ACCTCAGGTTGTGAAGAGCTCAAGAAAGACTCCAAACGGTTATTCATTAACATTACAAA	6100
OY	1187	GAAA 1190	
Db	6101	TACA 6104	

```

RESULT 15
US-09-949-016-16067/c
; Sequence 16067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16067
; LENGTH: 50217
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(50217)
;

```

Mon Feb 28 08:49:24 2005

OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16067

Query Match	8.1%;	Score 404.8;	DB 4;	Length 50217;
Best Local Similarity	76.2%;	Pred. No. 3.8e-84;		
Matches 550;	Conservative 0;	Mismatches 112;	Indels 60;	Gaps 3;

OY	349	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGACCGAGCTGCGTGTGCGCCAGGCTGGA	408
Dd	35211	TCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAAACGAGCTGCGTCTGTGCCCAAGCTGGA	35152
OY	409	GTCGAGTGGCGGATCTCGGCTCACTGCAGACTTCGCGCTCCGSGTTACAGCCATTCTCC	468
Dd	35151	GTGCAGTAGCGGAGATCTCAGCTCATCGAAGCTTCGCGCTCCGSGTTACAGCCATTCTCC	35092
OY	469	TGCCTCAGCCTCCCCAAGTAAGTGGGATACAGGCGCCGCTACTAAGCCCGGTAA'TTTT	528
Dd	35091	TGCCTCAGCCTCCCCAAGTAAGTGGGATACAGGCGCCGCTACTAAGCCCGGTAA'-TTT	35033
OY	529	TTTGTATTTTTTAGTAGAGAAGGGATTACCGTTTTTATGCGGGATGAGCTCGATTCTCTG	588
Dd	35032	TTTGTATTTTTTAGTAGAGAGGGGATTACCGTTTTTATGCGGATGAGCTCTTGA'TCTCTG	34973
OY	589	ACCCTGTGATCCGCGCGCTCGGCTCCCAAAGTGTGGGATTACAAGCGTGAAGCACCG	648
Dd	34972	ACCCTGTGATCCGCGCGCTCGGCTCCCAAAGTGTGGGATTACAAGCGTGAAGCACCG	34913
OY	649	CGCCCGGCGCATGATCATCTTCTTGAATACTGATGTGACAACTACTTAAGCATCAAA	708
Dd	34912	CGCCCGGCGCAAGAACTCTT-----	34893
OY	709	CTTACACCTTTAAATATGACATTTGGGCGAGGACCGGTGCTATGCTGTAA'TTCCAG	768
Dd	34892	-----AAAATCTAACAA'TAGGAGGCGAGTACGGTGTCTACGCTGTAA'TTCCAG	34841
OY	769	ACTTTGGAGGACAGAGTGGGTGAATCACTTGAAGCCAGAGTTTGAAGACAGCTTGCC	828
Dd	34840	AATTTGGAGGCTGACGCGGCGAGATCACTGAAGGTCAAGAGTTCAAAGACAGCTTGCC	34781
OY	829	AACATGTGAAC'TCTGTCTTTTACTTAAAAA'AAAAAAAAAAAAAAAAAATCAGCCGGT	888
Dd	34780	AATTTGCGAAACCCCCTTCTACT-----AAAAATCAAAAAA'TTAGCCAGGT	34732
OY	889	GTCGTGGGGACACCTGTATCCAGCTATGCTGAGGCTGAAGGACGAGAGCTTGA	948
Dd	34731	GTGTAAGTGGAGCTTATAGTCCAGCTACTCGGAGGCTGAAGGACGAAGAAT'TGCTTGA	34672
OY	949	ACCCTGAGGCGGAGTGTGAGTGGGCGGAGATCACTACCCGCTCCAGCTGGGCGA	1008
Dd	34671	ACCCAGGAAGAGAGCTTGCAGTGAAGCCGAGATCATGCCACTGCATCCAACTTAGCCA	34612
OY	1009	CAGAGCAAGACTCTGTCTCAATTAATTAATTAACAAAGAACAGCAAGCTTTGTGTACC	1068
Dd	34611	CAGAGCAAGACTCTCATCTCAAAAT'AAAAACA'AAANNNNNNNNNNNNNNNNNNNCAC	34552
OY	1069	TT'1070	
Dd	34551	TT'34550	

Search completed: February 27, 2005, 05:14:25
Job time : 797.948 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:12 : Search time 2691.32 Seconds
(without alignments)
11008.461 Million cell updates/sec

Title: US-09-936-271B-13_COPY_1_5000

Perfect score: 5000

Sequence: 1 gggccagagtgtaggaag.....tgatccacagtcgctgc 5000

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 19: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
- 21: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	9.4	525	US-10-029-386-13540	Sequence 13540, A
2	440.2	8.8	41454	US-10-087-192-1642	Sequence 1642, Ap
3	419	8.4	91352	US-10-300-611-4	Sequence 4, Appl1
4	418.4	8.4	115935	US-10-775-169-241	Sequence 241, Ap
5	417.6	8.4	26371	US-10-087-192-1450	Sequence 1450, Ap
6	417.6	8.4	91760	US-10-087-192-844	Sequence 844, App
7	417.6	8.4	136726	US-10-085-117-244	Sequence 244, App
8	415.2	8.3	17397	US-09-764-869-1945	Sequence 1945, Ap
9	415.2	8.3	17397	US-10-091-504-1945	Sequence 1945, Ap
10	415.2	8.3	17397	US-10-227-577-1945	Sequence 1945, Ap
11	415.2	8.3	19334	US-09-764-869-1943	Sequence 1943, Ap

12	415.2	8.3	19334	US-10-091-504-1943	Sequence 1943, Ap
13	415.2	8.3	19334	US-10-227-577-1943	Sequence 1943, Ap
14	415.2	8.3	19345	US-09-764-869-1944	Sequence 1944, Ap
15	415.2	8.3	19345	US-10-091-504-1944	Sequence 1944, Ap
16	415.2	8.3	19345	US-10-227-577-1944	Sequence 1944, Ap
17	415.2	8.3	29001	US-10-317-270-11	Sequence 11, Appl1
18	414.8	8.3	12970	US-09-764-891-7689	Sequence 7689, Ap
19	414.6	8.3	111084	US-10-723-860-1627	Sequence 1627, Ap
20	414.2	8.3	212231	US-10-087-192-1126	Sequence 1126, Ap
21	412.8	8.3	3294	US-10-027-633-115723	Sequence 115723, Ap
22	412.8	8.3	3294	US-10-027-633-115724	Sequence 115724, Ap
23	412.8	8.3	3294	US-10-027-633-115723	Sequence 115723, Ap
24	412.8	8.3	3294	US-10-027-633-115724	Sequence 115724, Ap
25	411.2	8.2	13170	US-10-741-600-15744	Sequence 5744, Ap
26	411.2	8.2	13170	US-10-741-600-17903	Sequence 17903, Ap
27	410.4	8.2	10445	US-09-764-891-6380	Sequence 6380, Ap
28	409.6	8.2	13328	US-10-741-600-17627	Sequence 17627, A
29	409.6	8.2	17303	US-10-741-600-17720	Sequence 17720, A
30	409.6	8.2	44801	US-10-741-600-17966	Sequence 17966, A
31	409.6	8.2	50000	US-10-364-505-6	Sequence 6, Appl1
32	409.6	8.2	50000	US-10-681-199-6	Sequence 681, Appl1
33	407.6	8.2	187844	US-10-719-993-6883	Sequence 6883, Ap
34	407.6	8.2	196686	US-10-087-192-484	Sequence 484, App
35	406.2	8.1	12970	US-09-764-891-7689	Sequence 7689, Ap
36	405.8	8.1	4433	US-09-764-891-6630	Sequence 6630, Ap
37	405.8	8.1	4433	US-10-091-572-668	Sequence 668, App
38	405.8	8.1	126001	US-10-175-492-13	Sequence 13, Appl1
39	405	8.1	32767	US-10-004-113-4	Sequence 4, Appl1
40	405	8.1	32767	US-10-394-948-4	Sequence 5704, Ap
41	405	8.1	67088	US-10-741-601-5704	Sequence 17804, Ap
42	405	8.1	67088	US-10-741-601-17804	Sequence 916, App
43	403.6	8.1	104068	US-10-741-600-1916	Sequence 1, Appl1
44	403.6	8.1	1503841	US-09-795-666-1	Sequence 1, Appl1
45	403.6	8.1	1503841	US-09-795-666-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-029-386-13540
; Sequence 13540, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13540
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 0.83
; OTHER INFORMATION: SWISSPROT HIT: Q9Y37, EVALU0.2.00e-34
; OTHER INFORMATION: NT HIT: AF135028.1, EVALU0.0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE388198.1, EVALU0.0.00e+00
US-10-029-386-13540

Query Match 9.4%; Score 472; DB 16; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.4e-102;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4529 TCCAGAAATCTGAGAGTCAGTTAGATTCACCTAACCCTGAGAGCTGAGCTGCTGTC 4588

Db	1	TCGAGGAATCTGTAGAGTCCAGTTGAGATCCAGTAAACCTTACCTGAGGCTGGGCTCTGTGC	60
Qy	4589	CTTGAAGCTTGAAGCTTGGGCTTTGAGAGGTGCCACTCTTAATTTCCAGGCGCTGCCCTTGC	4648
Db	61	CTTAGAGCTTGAAGCTTGGGCTTTGAGAGGTGCCACTCTTAATTTCCAGGCGCTGCCCTTGC	120
Qy	4649	CCCTCAGCATGTACAGACACCCACCCCTAGCTGTGCTGGGCTCTTGAAGTCCGAACCCAC	4708
Db	121	CCCTCAGCATGTACAGACACCCACCCCTAGCTGTGCTGGGCTCTTGAAGTCCGAACCCAC	180
Qy	4709	CCCCAGGCCAAGCCCCGCGCTCTGAGCCCGGCCCAACCATTTTCCGTTCCAGAGCATGT	4768
Db	181	CCCCAGGCCAAGCCCCGCGCTCTGAGCCCGGCCCAACCATTTTCCGTTCCAGAGCATGT	240
Qy	4769	TCTTGCCCAACAATGATGTTTCTCTGTGACCAACCTCTTAAACCGTGCCCTCTGGGAGCA	4828
Db	241	TCTTGCCCAACAATGATGTTTCTCTGTGACCAACCTCTTAAACCGTGCCCTCTGGGAGCA	300
Qy	4829	CCAGGACCTGGAGACTGGGGGCGGGGGAGAACGCGCCGGTCGGATTACACAGCAGCAGCGCAT	4888
Db	301	CCAGGACCTGGAGACTGGGGGCGGGGGAGAACGCGCCGGTCGGATTACACAGCAGCAGCGCAT	360
Qy	4889	CATCAATGGAATCCGACTTCGATTATGCAACCCAGCCGTCGTGGCAGGCGCGCTGTTGCTTAAG	4948
Db	361	CATCAATGGAATCCGACTTCGATTATGCAACCCAGCCGTCGTGGCAGGCGCGCTGTTGCTTAAG	420
Qy	4949	GCCCAACCAAGCTTACTTCGGGGGGGGGTGTGTGATTCACAGAGGCTGTCTC	5000
Db	421	GCCCAACCAAGCTTACTTCGGGGGGGGGTGTGTGATTCACAGAGGCTGTCTC	472

```

RESULT 2
US-10-087-192-1642
; Sequence 1642, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Engels, David W.
; APPLICANT: Moritz, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1642
; LENGTH: 41654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1642

```

	Query Match	8.8#	Score 440.2	DB 13	Length 41454
	Best Local Similarity	78.9%	Pred. No. 2,86-94		
	Matches 569	Conservative 0	Mismatches 133	Indels 19	Gaps 3
QY	345	TTCTTTT	TTTTTTTTTTTTTTTTTTTTTTTTTTGAACGGAGCTCGCTCTGCGCCACGGC	404	
Db	37225	TCCTTTGGCTCTTCTTTTTTTTTTTTTTTTTTTTGTGAACGGAGCTCGCTCTGTTGTCTACGGC	37228		
QY	405	TGAGATGACATGGCGGGATCTCGGCTCACTGCACAGCTCCGGCTCCGGGTTTACGCCATT	464		
Db	37285	TGAGATGACATGGCGGGATCTCGGCTCACTGCACAGCTCCGGCTCCGGGTTTACGCCATT	3734		
QY	465	CTCCGACCTACGCTCCCAAGTAAGTAACTGGGAATTACAGGCGCCGGCACATGACCGCGGTTAA	524		
Db	37345	CTCTGCTCAGCTCTCCCAATAGCTGGGAATTACAGGCACTCCGGCACATGACCGCGGTTAA	3740		
QY	525	T-----TTTTTGTATTTTAACTAGACAGGGGTTTCAACGTTTTTAAAGCCGGAGTGGCTCG	580		

Db	37405	TTTTATTATTTTAAATTTTAAAGTAAAGACGGGGTTTCACTGCTGTAAAGCAGATGGTTTCC	37466
Qy	581	ATCTCCTGACCTCGTATCCGCCCGCTTGGCTCCAAAGTCTGGGATTAACAGCGCTG	640
Db	37465	ATCTCCTGACCTCGTATCCGCCCGCTTGGCTCCAAAGTCTGGGATTAACAGCGCGG	37524
Qy	641	AGCCACCGGCCCCGGCCATGATCATTTCTTGAC-----TATGCTGATGTGACAACT	692
Db	37525	AGCCACCGGCCCCGGCCCCCTTGGTCTTAACTTCTTTCGCTTCAATTATTCAC	37584
Qy	693	ACCTAAAGCATCAGACTACACCTTAA-----ATATGCAGTTTGGGCGAGCACCG	745
Db	37585	ATATTAATTAATGACTTTGCTCTCTCTTCCAAATATGCAAGTTTAAAGAAATGAGGCGGCGCTCGG	37644
Qy	746	TGGCTCATGCTGTGTAATTCACGACTTTTGGAGGCGAGAGTGGTGTAATCACTTGAAGCC	805
Db	37645	TGGCTCATGCTGTATATCCAGACTTTTGGAGGCGAGAGCGAATCACTTGAAGTT	37704
Qy	806	AGGAGTTTGAACACAGCTTGGCCAAATGCTGAAACTCTGTCTTTATCTTAAAAA	865
Db	37705	AGGAGTTTGAACACAGCTTGGCCAAATGCTGAAACTCTGTCTTTATCTTAAAAA	37764
Qy	866	AAAAAAAAAAAAATTCAGCCGGGTTGCTGTGGGGCACTCTGTATCCAGCTATGTCTGAG	925
Db	37765	AAAAAACAATAATTAAGCAGGAGAGTGGCACTGCTGTATCTCCAGCTACTGGGAG	37824
Qy	926	GCTGAGGACGAGAGTCACTTGAACCTTGAAGCGGAGGTTGCACTGGGCGAGATCA	985
Db	37825	GCTGAGGACGAGAGTCACTTGAACCTTGAAGCGGAGGTTGCACTGGGCGAGATTC	37884
Qy	986	TCAACCGCCTTCAGCTCTGGGCGAGAGCAAGACTTGTCTCAATTAATTAATTAACA	1045
Db	37885	CCATTGCACTCAGCTCGATGACAGAGTGAATTCATCTCAAAAAA	37944
Qy	1046	A 1046	
Db	37945	A 37945	

```

RESULT 3
US-10-300-611-4
; Sequence 4, Application US/10300611
; Publication No. US20040097451A1
; GENERAL INFORMATION:
; APPLICANT: Ming-Yi Chiang
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: MODULATION OF NIDDOGEN EXPRESSION
; FILE REFERENCE: PTS-0059
; CURRENT APPLICATION NUMBER: US/10/300,611
; CURRENT FILING DATE: 2002-11-19
; NUMBER OF SEQ ID NOS: 136
; SEQ ID NO 4
; LENGTH: 91352
; TYPE: DNA
; ORGANISM: H. sapiens
US-10-300-611-4

```

Query Match	8.4%;	Score 419;	DB 17;	Length 91352;
Best Local Similarity	79.4%;	Pred: 4.5e-89;		
Matches 552;	Conservative 0;	Mismatches 125;	Indels 18;	Gaps 4;
Qy	350	TTTTTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGCTGTGCGCCAGGCTGAG	409	
Db	11334	TTATTATTATTATTATTATTATTATTTTGAATGATGATGCTGTGCTGCGCCAGTTGAG	11393	
Qy	410	TGAGTGGCGGGATCTCGGCTCACTGCAGCTCCGCTCCGGGGTTCAAGCAATTCCT	469	
Db	11394	TGAGTGGTGTGCTTGGCTTACTGCAAGCTCCAGCTCCCGGGTTCGCGCAATTCCT	11453	
Qy	470	GCTCAGCTCCCAATTAAGTGGGAATTAAGGCGCCCGGCAATGAGCCCGGTAATTTT	529	
Db	11454	GCTCAGCTCCCGAATTAAGTGGGAATTAAGCACTCCGACACGCGCCGGGTAATTTT	11512	


```

Query Match      8 4%; Score 418.4; DB 18; Length 115935;
Beet Local Similarity 70.2%; Pred. No. 6.8e-89;
Matches 597; Conservative 0; Mismatches 236; Indels 17; Gaps 2;

QY      355 TTTTTTTTTTTTTTTTTTTGAGAGGATCTCGCTCTGCCCGAGCTGGAGTGCAG 414
      |||||
Db       114273 TTTTTTTTTTTTTTTTTTTGAGGATCTCAGCTTTTCGCCAGCTGAGTGCAG 114332Z
      |||||

QY      415 TGGCGGATCTGGGCTCATGCAAGCTCCGCTCCGGGTTCAAGCAATCTCTGCTC 474
      |||||
Db       114333 TGGTGCAGTCTCTGCTCATGCAAGCTCTCTCCGGGTTCAAGCAATCTCTGCTC 114392Z
      |||||

QY      475 AGCTCCCAAGTAGCTGGGACTCAGGCGCGCGCACTAGCCCGGCTAATTTTTTGTGA 534
      |||||

```

```

RESULT 5
US-10-087-192-1450
; Sequence 1450, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITL OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1450
; LENGTH: 26371
; TYPE: DNA
; ORGANISM: Homo sapiens

```



```

; Sequence 1945, Application US/10091504
; Publication No. US2003005908A1
; GENERAL INFORMATION:
;   APPLICANT: Rosen et al.
;   TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;   FILE REFERENCE: PC007C1
;   CURRENT APPLICATION NUMBER: US/10/091,504
;   CURRENT FILING DATE: 2002-03-07
;   NUMBER OF SEQ ID NOS: 2442
;   Prior Application removed - See File Wrapper or Palm
;   SOFTWARE: PatentIn Ver. 2.0
;   SEQ ID NO 1945
;   LENGTH: 17397
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   US-10-091-504-1945

Query Match      8.3%; Score 415.2; DB 14; Length 17397;
Best Local Similarity 78.6%; Pred. No. 2e-88;
Matches 558; Conservative 0; Mismatches 123; Indels 29; Gaps 4;

Qy      349  TTTT+TTTTTTTTTTTTTTTTTTTTTTTTTGTGAGACGAGCTGCTGCTGTGCGCCAGCTGGA 408
Db      15150 TGTTCGTAACCTTTTTTGTGTGTTTAAGACAGACTTGTCTGTGCGCCAGGCTGGA 15209

Qy      409  GTGCAGTGGCGGGATCTCGGCTACCTGCAAGCTCCGCTCCGGGTTCAAGCATTCTCC 468
Db      15210 GTGACAGTGGCAGATCTCGGCTACCTGCAAGCTCCGCTCCGGGTTCAAGCATTCTCC 15269

Qy      469  TGCCTCAGCCCTCCCAAGTAGCTGGAGCTACAGAGCGCCCGCACCTACAGCCCGGCTAATTTT 528
Db      15270 TGCCTCAGCCCTCCCAAGTAGCTGGAGCTACAGAGCACCCGCGCACCAAGCCCGGCTAAT-TTT 15328

Qy      529  TTTGTATTTTTTAGTAGAGACGGGGTTTCAACGTTTTTAGCGGGATGGCTCGATCTCTG 588
Db      15329 TTTGTAATTTTAGTAGAGACGGGGTTTTTGCGGTTTAGCAGAGTGGTCTGATCTCTG 15388

Qy      589  ACCCTGTGATCCGCGCCCTCGGCTCCCAAAGTCTGGGATTACAGGCGCTGAGCCACG 648
Db      15389 ACCCTGTGATCTCGCGCTCGGCTCCCAAAGTCTGGGATTACAGGCGATGAGCTACG 15448

Qy      649  CGCCCGGCAGATCATCTTCTTGACATGCTGATGGAACAATCCTAAAGCATCAGA 708
Db      15449 CACCGGCGC-----GATGTTTGCACTCTTTAAACAGTAAACAAATCCCAAGCTTTTA-- 15501

Qy      709  CTCTAACCTTTAAATATGACAGTTTGGGCGAGGACCGGTGCTCATGCTGTATATCCAG 768
Db      15502 -----TAAATATATTAAGCTGGGGCTGGGCGGTGCTCATACCTGTATATCCAGC 15552

Qy      769  ACTTTGGAGGCGAGGTGGTGAATCATTTGAGGCCAGGAGTTTGAACACAGCTGGCC 828
Db      15553 AGTTTGGAGGCGCAGGCGAGGTGATCATCTTGAGCTCGGGTGTCAAGACCACTTGGCC 15612

Qy      829  AACATGGTGAACCTGTCTTTTACTAATAAAAAAAAAAAAAAAAAAAAAATGACCGGGGT 888
Db      15613 AACATGGTGAACCTGTCTTTTACT-----AAAAATCAAAAAATTTAGCCAGGT 15660

Qy      889  GTCTGGGCGCACCTGTAAATCCAGCTATGCTGAGAGCTGAGGACGAGAGTCACTTGA 948
Db      15661 GTGTGGGCGACACCTGTAAATCCAGCTATCTTGGAGAGCTGAGAGCAGAGAACACTTGA 15720

Qy      949  ACCCTGAGGCGGAGGTTGACGTGGGCGGAGATCATCAACGCGCTCCAGCTGGGCGA 1008
Db      15721 ACCCGGAGGCGGAGGTTGACGTGAGCGCGAGATTATGCCACTGCACCTGTGGGTGA 15780

Qy      1009  CAGGCAAGAGATCTGTCTCAAAATAAATTAATTAACAAGAACAGACAGT 1058
Db      15781 AAGGCAAGACTGTGTCAAAACATTAAAAATTAATTAATTAAGT 15830

RESULT 10
US-10-227-577-1945
; Sequence 1945, Application US/10227577

```

```

; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1945
; LENGTH: 17397
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-227-577-1945

Query Match      8.3%; Score 415.2; DB 17; Length 17397;
Beech Local Similarity 78.6%; Pred. No. 2e+88;
Matches 558; Conservative 0; Mismatches 123; Indels 29; Gaps 4;

QY      349      TTTTGTATTTTTTTAGTAGACGAGACTCGCTGTGCCTGCCCAAGCTGGA      408
DB      15150   TGTTTCGTAACCTTTTTTGTTGGTTGTTAGACAGAATCTTGCTGTGCCAAGCTGGA      15209

QY      409      GTGCAGTGGCGGAGTCTCGGCTCACATGCAAGCTCCGCGTTCAAGCATATTCTCC      468
DB      15210   GTGCAGTGGCAACGATCTCGGCTCACATGCAAGCTCCGCTCCGCGTTCAAGCATATTCTCC      15269

QY      469      TGCTCTAGCCTCCCAAGTACTTGGAGCTACAGGCGCGCCCACTAACGCCCGCTAATTTT      528
DB      15270   TGCTCTAGCCTCCCGAGTACTTGGAGCTACAGGCAACCGCCCACTAACCGCTAAT-TTT      15328

QY      529      TTGTATTTTTTGTAGTAGAGACGGGGTTTCAACCGTTTAAAGCGGAGATGGCTCGATCTCTG      588
DB      15329   TTGTATTTTTTGTAGTAGAGACGGGGTTTTCGCGTTTAAAGCAGAGATGCTCTCATCTCTCG      15388

QY      589      ACCTCGTATCGCGCCGCTCGGCTCCCAAAGTGTGGAGATTACAAGCGTGAAGCCACC      648
DB      15389   ACCTGTGATCTGCGCGCCGCTCGGCTCCCAAAGTGTGGAGATTACAAGCATGAGCTACTG      15448

QY      649      CGCCCGGCGCATGATCATCTTCTTGACTATGCTGATGTGACAGTACTTAAACCATCAGA      708
DB      15449   CACCCGCGC-----GATGTTGCACTCTTTAACTGTGAAAAAAATCCACACTTTTA--      15501

QY      709      CTCACACCTTAATAATATGCAAGTTTGGGACAGGACAGTGGCTCATGCTGTAAATTCACG      768
DB      15502   -----TAAATATATTAAGCTGGGCTGGGCGCGGTGCTCACACTGTATATCCAGC      15553

QY      769      ACTTTGGAGGCAAGAGTGGGTGAATCATCTTGAGGCAAGAGTTTGAACAGACCTGGCC      828
DB      15553   AGTTTGGAGGCAAGGACAGGTGATCATCTTGAGCTCGGGTGTTCAGAACCAAGCTTGGCC      15612

QY      829      AACATGCTGAACCTGTCTTTATCTAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGT      888

```


QY	889	GTGCTGGGGGCACACTGTAATCCCAAGTATGCTGAGAGCTGAGGCACGAGAGTCACCTTGA	948
Db	15654	GTGGTGGGACACACTGTAATCCCAAGTACTTGGAGAGCTGAGAGCAGAGAAATCACTTGA	15713
QY	949	ACCCGTGAGGGGAGGATTGTCAGTGGGGGGGAGATCAATGACGCGCCCTCCAGCCTGGGGCGA	1008
Db	15714	ACCCGGGAGGCGAGGGTGGCGTAGGCGGAGATTTATGCCACTGACTCCACTCTGGGGTGA	15773
QY	1009	CAGAGCAAGACTCTGTCTCAATTAATTAATTAACAAACGACAGACGAGT	1058
Db	15774	AAGGCGAAGACTCTGTCTCAAAACATTAAAAAAATTAATTAATTAATGAAG	15823

RESULT 13

```

1      ; Sequence 1943, Application US/10/227577
2      ; Publication No. US2004000575A1
3      ; GENERAL INFORMATION:
4      ; APPLICANT: Rosen et al.
5      ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
6      ; FILE REFERENCE: PC007C2
7      ; CURRENT APPLICATION NUMBER: US/10/227,577
8      ; CURRENT FILING DATE: 2002-08-26
9      ; PRIOR APPLICATION NUMBER: 10/091,504
10     ; PRIOR FILING DATE: 2002-03-07
11     ; PRIOR APPLICATION NUMBER: 09/764,869
12     ; PRIOR FILING DATE: 2001-01-17
13     ; PRIOR APPLICATION NUMBER: 60/1179,065
14     ; PRIOR FILING DATE: 2000-01-31
15     ; PRIOR APPLICATION NUMBER: 60/180,628
16     ; PRIOR FILING DATE: 2000-02-04
17     ; PRIOR APPLICATION NUMBER: 60/214,886
18     ; PRIOR FILING DATE: 2000-06-28
19     ; PRIOR APPLICATION NUMBER: 60/217,487
20     ; PRIOR FILING DATE: 2000-07-11
21     ; PRIOR APPLICATION NUMBER: 60/225,758
22     ; PRIOR FILING DATE: 2000-08-14
23     ; PRIOR APPLICATION NUMBER: 60/220,963
24     ; PRIOR FILING DATE: 2000-07-26
25     ; PRIOR APPLICATION NUMBER: 60/217,496
26     ; PRIOR FILING DATE: 2000-07-11
27     ; PRIOR APPLICATION NUMBER: 60/225,447
28     ; PRIOR FILING DATE: 2000-08-14
29     ; Remaining Prior Application data removed - See File Wrapper or PALM
30     ; NUMBER OF SEO ID NOS: 2442
31     ; SOFTWARE: PatentIn Ver. 2.0
32     ; SEO ID NO 1943
33     ; LENGTH: 19334
34     ; TYPE: DNA
35     ; ORGANISM: Homo sapiens
36     ; US-10-227-577-1943

```

Query Match	8.3%	Score 415.2	DB 17	Length 19334
Best Local Similarity	78.6%	Pred. No. 2.1e-88		
Matches 558; Conservative	0	Mismatches 123;	Indels 29;	Gaps 4;

QY	349	TTTTTTTTTTTTTTTTTTTTTTTTTTTGAACCGAGCTCGCTCTGTGCGCCAGGGCTGA	408
Db	15143	TGTTCTGTACTCTTTTTTGTGTGTGTAAACAAGACTCTGTCTCTGTGCGCCAGGGCTGA	15202
QY	409	GTGCAAGTGGCGGGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCAAGCCATTCTCC	468
Db	15203	GTGCAAGTGGCAAGATCTCGGCTCACTGCAAGCTCCGCTCCCGGGTTCAAGCCATTCTCC	15262
QY	469	TGCTCTAGAGCTCCCAAGTACGTGGACCTAAGGCGCGCCCACTACGCCCGGCTAAATTTT	528
Db	15263	TGCTCTAGAGCTCCCAAGTACGTGGACCTAAGGCGCGCCCACTAAGCCCGGCTAA-TTT	15321
QY	529	TTTTTATTTTTTAGTAGAGACGGGGTTTCAACCGTTTTAGCCGGGATGGCTTCATCTTCCG	588
Db	15322	TTTTTATTTTTTAGTAGAGACGGGGTTTTCGCGTTTAGCCAGGATGGCTTCATCTTCCG	15381

[illegible]

RESULT 14

```

US-09-764-869-1944
; Sequence.1944, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1944
; LENGTH: 19345
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-869-1944

```

Query Match	8.3%	Score 415.2;	DB 9;	Length 19345;
Best Local Similarity	78.6%;	Pred. No. 2.7e-88;		
Matches 558; Conservative	0;	Mismatches 123;	Indels 29;	Gaps 4;

QY	349	TTTATTTT	TTTTTTTT	TTTTTTTT	TTTTTTGA	GGGAGTCT	CGCTCTGCT	CTGCGCC	AGGCTGGA	408
QY	15151	TGTTCCGTA	CTCTTTTT	TGTTGTTG	TTAAGACAG	ACTCTTGC	TCTGTG	CGCCACAG	CGCTGGA	15210
Db	409	GTGCA	GTGGGGAT	CTCGGCTCA	CTGCAAGCT	CCGCGCTCC	CGGGTTCA	CGCCATTT	CTCC	468
QY	15211	GTGCA	GTGGGACAG	ATCTCGGCTCA	CTGCAAGCT	CCGCTCTCC	CGGGTTCA	CGCCATTT	CTCC	15270
Db	469	TGCTCA	GCTCCCCA	GTAGCTTGG	ACTAAGCG	CGCGGCACTA	CGCGCGCTA	TTT	528	
QY	15271	TGCTCA	GCTCCCCG	GTAGCTTGG	ACTAAGCA	CGCGGCACTA	CGCGCGCTA	TTT	15329	
Db	529	TTTGTAT	TTTTAGT	AGAGAGCG	GGGTTTCA	ACCGTTT	AGCGGGAT	TGACCTCAT	CTCCG	588
QY	15330	TTTGTAT	TTTTAGT	AGAGAGCG	GGGTTTTC	GCGTTT	AGCCAGAT	TGCTCAT	CTCTCCG	15389
Db	589	ACCTG	TGATCCG	CGCGCTCG	CGCTCCCA	AAAGTGT	GTGGAGTTA	CAGCGGT	GAGCACG	648

Db 1081 GGCTGAGTGCAGTGGCAGTATCTTGGCTCACTGCAAGCTCCGCTCTGGATTCAACC 1022
QY 462 ATTCTCTGCTGAGCTCCCAAGTAGTGGGACTACAGGCGCCCGCACTAGCCCGG 521
Db 1021 ATTCTCCGCTGAGCTCCGAGTAGTGGAGTCAAGGTGCGCCACTGCAACCAGC 962
QY 522 TAAATTT---TTTGATTTTATTTAGTAGAGAGCGGGGTTTCAACGTTTATAGCCGGATG 578
Db 961 TAAATTTTATTTTATTTTATTTATTTAGTAGAGAGCGGGGTTTCAACGTTTATAGCCAGATG 902
QY 579 CGATCTCTGAGCTCTGATCGCGCGCTGGCTCCCAAGTGTGGGATTTACAGCG 638
Db 901 CGATCTCTGAGCTCTGATCGCGCGCTGGCTCCCAAGTGTGGGATTTACAGCG 842
QY 639 TGAGCCACGCGCGCGCGCGAGTCAATCTTCTTGAATGATGAGCAATGACTTAA 698
Db 841 TGAGCCACGCTGCTGCGC-----TGTTTCTGATCTTGA 808
QY 699 AGGCATTCAGACTTACCTTTTAAATATGAGTTTGGCCAGGACCGTGGCTCATGCT 758
Db 807 AAACATTCAGCTTTTCACTAAATAATATATTTGGCCAGATGCGGTGCTTATGCTG 748
QY 759 TAAATTCAGACTTTTGGAGGCGAGAGTGGGTGATCACTTGAAGCCAGAGATTGAGAC 818
Db 747 TAAATTCAGACTTTTGGAGGCGTGAAGTGGGTGATCAC--GAGGTCAAGAGTTTGAAGC 690
QY 819 CAGCTGGGCGCAAGTGGTGAACCTGTCTTACTTAAATTAATTAATTAATTAATTAATTA 878
Db 689 CAGCTGGGCGCAAGTGGTGAACCTGTCTTACT-----AAAAATTAATTAATTAATTA 642
QY 879 TCAACCGGCTGTGTGTGGGCGACACTGTATATCCAGTATGTCAGAGCTGAGCGACGAG 938
Db 641 TTAGCTGTGTGTGTGGGCGGCGCTGTATATCCAGTATCTTGGAGGCTGAGGCGAGAG 582
QY 939 AGTCACTTGAACCTTGGAGGCGGAGTGTGCAATGAGGCGAGATCAATCAACCCCTTCCA 998
Db 581 AATCGCTTGAACCGAGAGGCGGAGTGTGCAATGAGGCGAGATGAGGCGAGTGTGCACTCCA 522
QY 999 GCTTGGGCGAGAGGAGCACTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTA 1052
Db 521 GTCTGGGCAAGAGTGAAGCACTGTCTCGAAAAATTAATTAATTAATTAATTAATTA 468

RESULT 2
AF101960 2971 bp DNA linear GSS 06-NOV-2000
LOCUS AF101960 Human Homo sapiens genomic clone PTMB59.14, genomic survey
DEFINITION
ACCESSION AF101960
VERSION AF101960
KEYWORDS AF101960.1 GI:4193786
SOURCE GSS.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Begler, G., O'Brian, K.C., Kim, Y.C., Schreiber, G. and Pittlerle, D.M.
TITLE A 1.4-Mb high-resolution physical map and contig of chromosome
segment 11p15.5 and genes in the LOH1A metastasis suppressor
region
JOURNAL Genomics 55 (2), 164-175 (1999)
MEDLINE 99134294
PUBMED 9933563
COMMENT
Contact: Begler G
Medicine and Radiology
Duke University Medical Center
Box 2610, MSRB, Room 117, Durham, NC 27710, USA
part of a 1.4 megabase contig including the LOH1A metastasis
suppressor region Bin T
Class: unknown.
FEATURES
Source Location/Qualifiers
1..2971
/organism="Homo sapiens"

ORIGIN
Query Match 7.5%; Score 373.8; DB 8; Length 2971;
Best Local Similarity 71.2%; Pred. No. 6.7e-23;
Matches 566; Conservative 0; Mismatches 192; Indels 37; Gaps 4;
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="11p15.5"
/clone="PTMB59.14"
/clone_11b="Human"

QY 339 ATCATCTCTTT 398
Db 456 ATAAATTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 515
QY 399 CCAAGCTGAGTGCAGTGGCGGAGTCTGGCTCACTGCAAGTCCGCTCCCGGGTTCA 458
Db 516 CCAAGCTGAGTGCAGTGGCGGAGTCTGGCTCACTGCAAGTCCGCTCCCGGGTTCA 575
QY 459 GCCATTCTCTGCTCAGGCTCCCAAGTAGTGGAGTCAAGAGCCCGCCACTAGCC 518
Db 576 GCGATTCTCTGCTCAGGCTCCCAAGTAGTGGAGTCAAGAGCCCGCCACTAGCC 635
QY 519 GCGTAAATTTTTTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 578
Db 636 AGCTAA--TTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 693
QY 579 CGATCTCTGAGCTCTGATATCCGCGCGCTCGGCTCCCAAGTGTGGGATTTACAGCG 638
Db 694 CAAATCTCTGAGCTCTGATATCCCACTCAGCTCCCAAGTGTGGGATTTACAGCG 753
QY 639 TGAGCCACGCGCGCGCGAGTATCTTCTTGAATGATGATGATGATGATGATGATGATGATGAT 698
Db 754 TAAAGCAGCGACTGCGC-----CAACTATATTTTAA 787
QY 699 AGGCATTCAGACTTACCTTTTAAATATGAGTTTGGCCAGGACCGTGGCTCATGCTG 758
Db 788 TGTGTAAGACTGAATGTTTCTTAAGATCAAGAACAGGCGAGAGTGTGCTGAGCGCTG 847
QY 759 TAAATTCAGACTTTTGGAGGCGGAGTGGGTGATCACTTGAAGCCAGAGATTGAGAC 818
Db 848 TAAATTCAGACTTTTGGAGGCGGAGTGGGTGATCACTTGAAGCCAGAGATTGAGAC 905
QY 819 CAGCTGGCGCAAGTGTGAACCTGTCTTACTTAAATTAATTAATTAATTAATTAATTAATTA 878
Db 906 CATCTCGGTAAACAGGTGAACCCGCTTCTACT-----AAAAATTAATTAATTAATTA 958
QY 879 TCAACCGGCTGTGTGTGGGCGACACTGTATATCCAGTATGCTGAGGCTGAGCGACGAG 938
Db 959 TTAGCTGTGTGTGTGGGCGACCTGTATGCTCCAGCTACTCGGAGGCTGAGCGCGGAG 1018
QY 939 AGTCACTTGAACCTTGGAGGCGGAGTGTGCAATGAGGCGAGATCAATCAACGCCCTTCCA 998
Db 1019 AATGTGTGAACCCAGAGGCGGAGTGTGCAATGAGGCGAGATCAATCAACGCCCTTCCA 1078
QY 999 GCTTGGGCGAGAGCAAGTCTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1058
Db 1079 GCTTGGGCGAGAGGCGAGATCTCGACTCAAAAAAAGATTTAGAAATGAAGACAG 1138
QY 1059 TTGTGTACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1118
Db 1139 GATATCTGCTACACTTCTTCTTCAACACTGTACTGAATTAATTAATTAATTAATTAATTA 1198
QY 1119 ATTAAGAAATTAATTA 1133
Db 1199 GGCAGAAAAAGAAA 1213

RESULT 3
CRS92231 1715 bp mRNA linear HNC 21-UTL-2004
LOCUS CRS92231
DEFINITION full-length cDNA clone GSD1023YE11 of Placenta Cot 25-normalized
of Homo sapiens (human).

ACCESSION	CR592231
VERSION	CR592231.1 GI:50473038
KEYWORDS	HTC, CNSLT_cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : <a homo="" href="http://fulllength.invitrogen.com/InvitrogenCorporation1600ParadeyAvenue2(bases1to1715)Genoscope.DirectSubmissionSubmitted(20-JUN-2004)Genoscope-PRANCE(E-mail:secref@genoscope.cns.frBP191.91006.EVRY.cedex-PRANCE(E-mail:secref@genoscope.cns.frWeb:www.genoscope.cns.fr)1ststrandcDNAwasprimedwithaNciI-oligo(4nt)primer.Fiveprimariesenriched,double-strandcDNAwasdigestedwithNciIandclonedintotheNotIandBclI sites of the pCMVSPORT6vector.Librarywasnormalized.LibrarywasconstructedbyLifeTechnologies,a division of Invitrogen</td></tr><tr><td>FEATURES</td><td>Location/Qualifiers</td></tr><tr><td>source</td><td>1..1715</td></tr><tr><td></td><td>/organism=" sapiens"<="" td="">
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CSOD1023YR11"
	/cissue_type="Placenta Cot 25-normalized"
	/plasmid="pCMVSPORT_6"
ORIGIN	
Query Match	7.3% Score 365; DB 3; Length 1715;
Best Local Similarity	71.5%; Pred No.5.1e-22;
Matches 551; Conservative	0; Mismatches 135; Indels 85; Gaps 2;
355	TTTTTTTTTTTTTTTTTTTTTGTGAGACGAGTCTGCTCTGTGCGCCAGGCTGTGAGTGCAG 414
DB	701 TTTTTTTTTTTTTTTTTTTTGTGACGACAGTCTTGTCTGTGCGCCAGGCTGTGAGTGCAG 760
415	TGGCGGAGTCTGCGCTCACTGCAGACTCGGCTCCCGGGTTACGCCCATTTCTCTGCTCTC 474
DB	761 TGGCCCGAGTCTGCGCTCACTGCAGACTCTGTCTTGTGGGTTTCATCCATTCTCTGCTCTC 820
475	AGCCTCCCAAGTACGTGGGACATACAGGGCCGCGACATACGCGCGGCTAAATTTTTTGTGA 534
DB	821 AGCCTCCCAAGTACGTGGGACATACAGGGCCGCGACATACGCGCGGCTAAATTTTTTGTGA 880
535	TTTTTAGTAGAGACGGGGTTTCAACGGTTTTAGCCGGGATGAGCCTGCATCTCCTGCACCTG 594
DB	881 TTTTAGTAGAGATGAGGGGTTTCAACGGTTTGTAGCCATGATGATGATCTCTCTGACCTTG 940
595	TGATCCGCGCCGCTTGCGCTCCCAAGTGTGGGATTAACAGCGGTGAGCCACCGGCGCCG 654
DB	941 TGACTGTGCTGTGCGCGCCCGCCCAAGGGCTGGATTAACAGCGGTGAGCCACCGGACCCA 1000
655	GCCA----- 658
DB	1001 GCGACTTTTCTTGATTAAGAGACCATGACTATGAGGGCTGGTTCTGGCCAGTTTATTAAG 1060
659	-----TGATCATCTTCTTGACTATGCTGATGTGACAGATCTAAAGCC 702
DB	1061 GCTAAGACACCTTGTCTCTCTCTGCTGTCTGACAGAGCATTTTACTTTAAGGTCTAAATG 1120
703	ATCAGACTTACCTTTAAATATGAGATTGGGCGACGACCGTGGCTCATGCTGTAT 762
DB	1121 TTATCATTTAAGTATTAAGTCTCACACCTTAGGCTGTGGGCTGTGGCTCAGCGCTGTAT 1180
763	TCCACACATTTTGGGAGCGACAGGTGGGTGAATCTTAGAGCCAGAGATTGAGCCAGC 822

Df	1181	CCCAAGCACTTTGGGAGCGTGAAGGTGGATGACTTGAGGTTCAGAGATTGGACACCAGC	1240
Oy	823	CTGGCCAAACATGTGTGAACCTCTGTCTTACTTAATAAAAAAAAAAAAAAAAAAATTCAG	882
Df	1241	CTGGCCAAACATGTGTGAACCTCTGTCTTACTTAATAAAAAAAAAAAAAAAAAAATTCAG	1287
Oy	883	CCGGGTGTGCGGGGGCACACCTGTATATCCAGCATCTGAGAGCTGAGCAGAGATC	942
Df	1288	CTGAGTGTGTGTGTGCACGCTGTATATCCAGCATCTTGGAGACTGAGAGAGAATC	1347
Oy	943	ACTTGAACCTTGAGGCGGAGTTGCACTGTGGCGGAGATCAATCACCCTCCAGCTT	1002
Df	1348	ATTGGAACCTTGAGGCGGAGATGATGAGTGTGAGATGGCGCCACTGCACCTCAGTCT	1407
Oy	1003	GCGGACAGACGAAGACTCTGTCTCAATAATAATAATAAACAAACAACA	1053
Df	1408	GGGTGACAGATTAAGACTCCGTCGCGAAAAACAAAACAATAACA	1458
RESULT 4			
HSM805366/c			
LOCUS	HSN805366	2429 bp.	mRNA linear HTC 22-SEP-2004
DEFINITION	Homo sapiens mRNA; CDNA DKFZP586G2417 (from clone DKFZP586G2417).		
ACCESSION	ALB34319		
VERSION	ALB34319.1	GI:21739931	
KEYWORDS	HTC.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
AUTHORS	1 (bases 1 to 2429) Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fodor,G., Han,M. and Wiemann,S. The German CDNA Consortium Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberberg, GERMANY		
CONSRMT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkz-heidelberg.de; Sequenced by BMFZ (Biomedical Research Center at the Heinrich-Heine-University, Duesseldorf/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZ586G2417) is available at the RZPD Deutsches Reisourcezentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZ586G2417 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.		
JOURNAL	Location/Qualifiers		
COMMENT	1..2429		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="RZPD:DKFZP586G2417"		
	/db_xref="taxon:9606"		
	/clone="DKFZP586G2417"		
	/tissue_type="uterus"		
	/clone_id="586 (synonym: hute1). Vector pSport1; host		
	DH10B; sites NotI + SalI/MluI"		
	/dev_stage="adult"		
	/note="hypothetical protein"		
	1..2429		
	/gene="DKFZP586G2417"		
	308..1024		
	/gene="DKFZP586G2417"		
	/codon_start=1		
	/product="hypothetical protein"		
	/protein_id="CAD8988.1"		
	/db_xref="GI:21739932"		
	/db_xref="UniProt/TREMBL:Q8ND94"		
	/translation="MLSPCLMLAVTFVLPAPQAIPADPFEBEADDETETAMPPLP		
	HWCPDPYVHLGVCKELRGVPACITGLSSPAQPPPRNGIVRIABEGRAVV		
	AWCAPRSYVAILWILLMDSEAKOKGPINAATVRRELKGLKKGGIVVVCVAENAG		
	ASRPQAGSGGLEGADIPIARGPCSRILAVPNPTLVHANVGVTALLALSICALVWHF		
FEATURES			
SOURCE			
gene			
CDS			

LOCUS	BC033224/c	1641 bp	mRNA	linear	HTC 27-JUN-2002
DEFINITION	BC033224	Homo sapiens, clone IMAGE:5022014,	mRNA.		
ACCESSION	BC033224				
VERSION	BC033224.1	GI:2161936			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 1641)				
TITLE	Strauberg, R.				
JOURNAL	Direct Submission				
	Submitted (25-JUN-2002) National Institutes of Health, Mammalian				

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

CDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNI).
DNA dequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fiehl, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Kryzhanek, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Parbhoo, Parvaneh Saeeadi, Jacqueline
Schell, Duane Smalins, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasha van den Bosch, Jill Vardy,
George Yang, Scott Zuydam.

This clone has the following problem: frame shifted.

```
Location/Qualifiers
1. 1641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5022014"
/tissue_type="muscle, rhabdomyosarcoma"
/clone_id="NH MGC 17"
/lab_host="DH10B-R"
/ncbi="Vector: pOTB7"
```

ORIGIN

	Query Match	6.7%	Score 337.2	DB 3	Length 1641
	Best Local Similarity	75.3%	Pred. 1.1e-19		
	Matches	500	Conservative	0	Mismatches 120; Indels 36; Gaps 5;
QY	348	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGCTCTGTGCGCCACAGCTGG	407		
DB	1118	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGCTTGCTCTGTGCATCCAGGCTGT	1059		
QY	408	AGTGCAGTGGCGGGATCTCGGCTCACTGCAGAGCTCCGCTCCCGGGTTCACGCCATTCTC	467		
DB	1058	AGTGCAGTGGCGGTGATCTCACCTCACTCAAGCTCTGCTCTCTCGGGTTCAGGCCATTCTC	999		
QY	468	CTGCCTCAGCTCCCAAGTAGCTGGGATTAACAGGGCGCCGCGCATAGCGCGGCAATT	527		
DB	998	CTGCCTCAGCTCCCAAGTAGCTGGGATTAACAGGGCGCCACACACACCCGACTTAA-TT	940		
QY	528	TTTTGTATTTTATGACAGCGGGTTCAACCGTTTATGCGGGATGGCGCTTCGATCTCT	587		

Db	933	TTTTGTATATTTTCACTAAGACAGAGGTTTCACTTGTATTAGCCAGATAGTCTCCATCTCTCT	880
OY	588	GACCTCGTATCGGCGCCGCTCGGCTTCCCAAGTGGTGGATTACAGCGGTGAGCCACC	647
Db	879	GACCTCATATGATCCACTGGCTCGGCTTCCCAAGTGGTGGATTACAGGATGAGCCACC	820
OY	648	GCGCCCGGCGCATGATCATCTTCTTGAATATGTCTGATCTGACCAATGACTTAAAGCATACAG	707
Db	819	ATGCCCCAGCCCTAAACTCTCATTTCTTAATTGGACAATTTTAAAGATATTTTAACTCCAGA-	761
OY	708	ACTCTACCTTTTAAATATGACAGTTTGGGCCAGGACCGTGGCTATGCTGTAAATTCAG	767
Db	760	-----CTATATAAAAAAGAGCTTACAGCCAGGACAGCATGCTGTCACCATGTAAATTCACAG	709
OY	768	CACTTTGGGAGGCGAGAGTGGGTAAATCACTTGAAGCCAGAGATTATAGACCAAGCTGGC	827
Db	708	CACTTTGGGAGGCGGAGGT-AGTGTATCATCTTGAAGCCAGGAGTTCAAGACCAAGCTGGG	650
OY	828	CAACATGATGTAACCTGTCTTTTACTTAAAAAATTTAAAAAATTTAAAAAATATCAGCGGG	887
Db	649	CAACAGAGCCGAGACTCAAACTCTACGAAAAATTTAAAAAATTA-----GCGGG	602
OY	888	TGTGTGGGGGACACCTGTAAATCCACACTATGCTGGAGGCTGAGGACAGAGTCACTTG	947
Db	601	ACGTGGGGCACCCCTGTGTAGTTTCAACTACTCAGGAGG-----ATTGCTTG	555
OY	948	AACCTTGGAGGCGGAGGTTGACAGTGGGCGAGATCAATCAACCGCCCTCCAGGCTGGAGCG	1007
Db	554	AACTGGGAGATTCAAGGTTGAGAGCCGAGACATAGCCACTCTCCAGCTGGAGTG	495
OY	1008	ACAG 1011	
Db	494	ACAG 491	

RESULT	7
LOCUS	BC038630/c
DEFINITION	BC038630 2772 bp mRNA linear HTC 04-MAR-2003
ACCESSION	Homo sapiens, similar to hypothetical protein FLJ20489, clone IMAGE:5263792, mRNA.
VERSION	BC038630
KEYWORDS	BC038630.1 GI:24116283
SOURCE	HTC.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. Strausberg,R.
TITLE	Direct Submission
JOURNAL	Submitted (15-OCT-2002) National Institutes of Health, Mammalian

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

Tissue Procurement: Mikielos Palkovits, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
 Toshiyuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
 contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahey, Erin Helton, Mark Kerteman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>
 Series: IRAK Family: 73 Row: 1 Column: 4
 This clone has the following problem: retained intron.

FEATURES
 Location/Qualifiers
 1..2772

Db	241	CTCTGACCTGCTATCTGCCCCGCTTCCACCTCCCAAGTCTGGGATTAACAGCGCTGAG	300
QY	643	CCACCGCGCCCGGCCATGATCATCTCTTGACTATGCTGATGTGCAAGTACTTAAAGCC	702
Db	301	CCACCATGCTGGGCCATTATATATTTCTTTTCTTTGTCGTTAATTGAGCAATGACCA	360
QY	703	ATCAGACTCTTACCTTTAAATATGAGTTTGGGCGCAGGCCCGTGGCTCATGCTGTAT	762
Db	361	GAGGACTTGACGAGACTACT-----GTTGGCCGGCGTGTGACTCATATCTGAAT	414
QY	763	TTCAGCATTTTGGAGGAGAGATGGGTGAATCACTTAGGCGCAGAGATTGAGCCAGC	822
Db	415	CCGAGTACTTCAGGAGAGCCAGAGAGGTGGAATCTTTGAGCCGGAATTTGAGACCAAG	474
QY	823	CTGGCCAAATATGTGAAACTGTCTTTTACTTAAAAAATTTTTTTTTTTTTTTTAAATCAG	882
Db	475	CTGGGCACAATGCGCAAAACCTGTCTCTACAAAAATTA-----	513
QY	883	CCGGGTGTGTTGGGGCACACCTGTAAATCCCAAGTATGCTGAGGCTGAGGACAGAGT	942
Db	514	---GGTGTGTGTCCACACGCTGTGTGTCCAGCTACTGGAGGCTGAGTAAATGCTC	570
QY	943	ACTTGAACCTTGAGGCGGAGGTTGCACTGGGCGCAGATCATCAATCCGCTCAGGCT	1002
Db	571	ACTGAACCTTGAGGTTAAAGGCTGCACTGAGCGGAGATCTCGCACATGTACTCAGGCT	630
QY	1003	GGGCGACAGACG-AGATCTGTCTCAATAATTAATTAATTAACCAACGAAACGACGTTG	1061
Db	631	GAGTGGCAGACCAACCTATCTCANAAATATATTTANNAAGAAAAAATTTTTAAAAA	690
QY	1062	TTGTACC 1068	
Db	691	TCGTGCC 697	

RESULT 9

LOCUS AB012089

DEFINITION Homo sapiens gene, genomic survey sequence.

ACCESSION AB012089

VERSION AB012089.1 GI:3149956

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE AUTHORS Kurahashi, H., Sakamoto, M., Ono, J., Honda, A., Okada, S. and Nakamura, Y.

TITLE Molecular cloning of the chromosomal breakpoint in the LIS1 gene of a patient with isolated lissencephaly and balanced t(8;17)

Journal Hum. Genet. (1998) in press

Journal Title 2 (bases 1 to 5528)

Journal Authors Kurahashi, H.

Journal Direct Submission

Journal Submitted (09-MAR-1998) Hiroki Kurahashi, Biomedical Research Center, Osaka University Medical School, Division of Clinical Genetics, Department of Medical Genetics, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail: kura@cigene.med.osaka-u.ac.jp, Tel: 81-6-879-3381, Fax: 81-6-879-3389)

Features

source

1..5528

location/Qualifiers

organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17p13.3"

QY	3299	AAAGAACGAGCGCGGGCGGGGGGCTGACAGCGCTGTAATCCGAGCCCTTTGGAGGCGCGAG	3358
Db	2123	AAATATACGAGCGAGCGCGGGGCTGACAGCGCTGTAATCCGAGCACTTGGAGGCGCGAG	2182
QY	3359	CAGCGGATTCACCTGAGGTCGGAGAGATCGAGCCAGCGCTGACCAATGAG-AACCCG	3417
Db	2183	CAGGTGATTCACCTGAGGTCAGAGGTTTCGAGACCAGCGCTGGCTAAATGTAACCCG	2242
QY	3418	TCTTACTTAAATATCAAAAAATTAGCTGGTGTGTGGCGCGCTGTATATCCAGCTA	3477
Db	2243	TCTTACTTAAATATCAAAAAATTAGCTGGAGGTGTGTGATATCCTGTATATCCAGCTA	2302
QY	3478	CTCGGAGACCTGGGCGAGAAAAATCGCTGAAACCCGGAGGTGAGGTTGCGGTGAGCG	3537
Db	2303	CTTGGAGAGCTGAGGCGAGAGAAATTGCTGAACCCAGAGGCAAGAAATTGCAGTGATTTG	2362
QY	3538	AGATCGCGCCATTGCACTCCAGCGCTGGGCAACAGAGGAAATCCGCTTCGAAAGAAA	3597
Db	2363	AGATTGCAACCACTGATCTCAGCGCTGGGCAAAAAGAGAAAATTGTGTCCAGGAAA	2422
QY	3598	AAAAA-----AAAAAAGGTAAAGAACAGTGAATGGC	3633
Db	2423	AAAAAAGTATATATATATATATACACACACATTTTGTATTTATCAATTTCAATAGA	2482
QY	3634	ACGGAGGACATGATGATGAGTGGGGGACATGATGATGATGATGATGATGATGAGG	3693
Db	2483	AGGGGAGAAATCTTGATTTGACATTTGGGTAATTTGAATTTTCTCCAGATCTTTTTTTT	2542
QY	3694	AGGAGATTGACAGGATTGAGAGGACATGTTTTCATCTGAGAAATTCAGAAACCTAGCGCTG	3753
Db	2543	TTTTTCTGAGACAGGAGCTCTCTCTGCGCTGACATGAGTGCAGTGACAGATCTTG	2602
QY	3754	CTTTTCCCTTCATGTGGCCCCCTTAAGCTGAGCCCTCT-----TTCTGTGCTCGC	3805
Db	2603	GCTTATGCAAGCTTCACCTCCAGGTTTCAGTGTCTTGTGCTGACGCTCTTGATGATCT	2662
QY	3806	TTTTGGAACCTTACCTCGCGCCATGAGCTGAGCCCACTCTTCTCAACACGCGCC	3865
Db	2663	GGGATTAACGAGCATGCGGACACACGCGCGGTATTTTGTATTTTAAGTAGAGATGGG	2722
QY	3866	CTAGGCGACATCTAGTGTGAGCCCGGCTTAGGCGCACCCCTT-----TGGGCAGGCTG	3920
Db	2723	TTTCAACCATGTGGTCAAGACTGTATCTCAAACTCTGACCCCGTATGCTCCACCTTGGC	2782
QY	3921	CACCCCATTTCTGTGGGTACTTCTTGAACCCCTTCAAAAGTACAGACTTTTTTTTT	3980
Db	2783	CTCCCAAGTCTGGGATTTACAGGCGTGTAGGCATATCGCGCTGGGCAAGTATATTTTATTT	2842
QY	3981	TTTTTTTGGAGACAGTCTTGTCTCTCTCTCCAGGCTGGAGTGTGAGTGTGATCTCGGC	4040
Db	2843	TCTTTTTTGAAGACGAGATGTAGAGTCT--TGTGCCAAGAGTGCATATGGCGTATCTTGGC	2901
QY	4041	TCACTGCACCTTGTGCTCTCCAGGTTAAGTATTTCTGTGCTTCACCTCTCTGATGAC	4100
Db	2902	TCACTGCACCTTCACCTCCAGGTTTCAAGAGATTTTCTGTGCTTCACCTCTCCAGTATGC	2961
QY	4101	TGGGATTAACAGTGTGGGCGGCAACAGCGCTGGCTAATTTTGTGTCTTATGAGAGACAG	4160
Db	2962	TGGGATTAACAGTGTGGGCGGCAACAGCGCTGGCTAATTTTGTATTTTATGAGAGATGGA	3021
QY	4161	GTTTTACCTTGTGGGCGAGGCTGTCTCAAACTCCCAACCTCAGATATCCGCCACCTC	4220
Db	3022	GTTTTACCAATTTGGCGAGGCTGTGTCTTCCAACTCCTCAGTCAAGTATATCCCGGCTC	3081
QY	4221	GGCTTCCAGAGTGTGGGTTTACAGCGTGTAGGCAACCGCCCAAGCCCAAGTTC	4275
Db	3082	GGCTTCCCAAGTGTGTAGGTTTACAGGTGTGTAGGCAACCGCACTCTGTGTGATTC	3136

RESULT 11	652 bp	mRNA	linear	EST 01-AUG-2002
LOCUS	AUI20416	HEMBB1	Homo sapiens cDNA clone HEMMB1000637 5', mRNA	
DEFINITION	AUI20416	HEMBB1	Homo sapiens cDNA clone HEMMB1000637 5', mRNA	
ACCESSION	AUI20416	HEMBB1	Homo sapiens cDNA clone HEMMB1000637 5', mRNA	
VERSION	AUI20416	HEMBB1	Homo sapiens cDNA clone HEMMB1000637 5', mRNA	
KEYWORDS	AUI20416.1	GI:10935651	EST.	
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 652)			
	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Negai, T., Sugano, S. and			
	Isozaki, T.			
TITLE	HRI human cDNA project			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Takao Isozaki			
	Genomics Laboratory			
	Helix Research Institute			
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan			
	Tel: 81-438-52-3975			
	Fax: 81-438-52-3986			
	Email: genomcs@hri.co.jp			
	HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.			
FEATURES	Location/Qualifiers			
source	1..652			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/clone="HEMBB1000637"			
	/issue_type="whole embryo, mainly body"			
	/dev_stage="embryo, 10 weeks"			
	/clone_lib="HEMBB1"			
	/note="Vector: pME18SFL3"			
ORIGIN				
Query Match	6.5%; Score 325.2; DB 1; Length 652;			
Best Local Similarity	74.7%; Pred. No. 1, 9e-18;			
Matches	489; Conservative 0; Mismatches 148; Indels 18; Gaps 6;			
352	TTTTTTTTTTTTTTTTTTTTTTTGTGAGACGAGCTCTGCTC-TGTGCCACGAGCTGAGT	410		
3	TCTTTCAACAATTTTTTTTTTTTTTTTGTGAGACGAGATTTTGTCTTGCCGCTCAGAGCTGGAGT	62		
411	GCAATGCGGGGATTTGCGCTCATCTGCAAGCTCCGCGCTCCGGGGTTACAGCATTTCCCTG	470		
63	GCAATGCGGGGATTTGCGCTCATCTGCAAGCTCCGCGCTCCGGGGTTACAGCATTTCCCTG	122		
471	CCTAGGCTCCCAAGTAGTGTGGGACATACAGAGCGCGCGCCGACATACGCGGCTAATTTTTT	530		
123	CCTAGGCTCCCAAGTAGTGTGGGACATACAGAGCGCGCGCCGACATACGCGGCTAATTTTTT	180		
531	TGTAATTTTGTAGTAGACGCGGGTTTCAACGCTTTTGTAGCGCGGATGAGCTCGATCTCTGAC	590		
181	TGTAATTTTGTAGTAGAGTAGGGTTTCAACATGTTGTGGCAAGGCTGTCTCAAACTCTGAC	240		
591	CTC--GTGATCGCGCGCGCTCTGCGCTCCCAAAAGTGTGGGATTTACAGGCGTGAAGCCACG	648		
241	CTCAGGTGATCCCACTGCTCGCGCTTCCCAAGGTGCGGATTTACAGGCGTGAAGCCACG	300		
649	CGCGCGGACATGATCATCTCTTGAACATGCTGATGATGATGATGATGATGATGATGATGATGAT	708		
301	TGCTGTGCTTTTATTTTATTTTGT	360		
709	CTCTACACCTTTAAATATGACATTTGGGCGACGACCGTGGCTGATGCTGTAAATTCAGC	768		

Db 361 ACAAAATATATA-----TACGCCGGGTCTGTGGCTCAGCCTGTAATCCAGC 411
Qy 769 ACTTTGGAGGAGAGGTGGTGAATCATCTTAGAGCCAGATTGAGACCAAGCTG3CC 828
Db 412 ACTTTGGAGGAGAGGTGGTGAATCAC--GAGATCAGGAGATCGAGACCATCTG3CT 469
Qy 829 AACATGTTGAACCTCTCTTACTATAAAAAAAAAAAAAAAAAAATCAGCCGGT 888
Db 470 AACACATGAAACCTCTCTCTACTTAAAAAAAAAAAAAAAAAAG--TACCTG3CC 527
Qy 889 GTCGTGGGAGACACTTAAATCCAGCTATGCTGAGGCTGAGCAGAGATGACTTGA 948
Db 528 ATGTGTGCTGCACTGTAAGTCCAGCTACTCAGAGAGCTGAGCAGAGATGCTTGA 587
Qy 949 ACCCTGAGGCGGAGGTTCAGTGGCCGAGATCATCACCCTCCAGCCTG 1003
Db 588 ACCGGGAGGAGAGGTTCAGTGGCTGAGAGCAGCAGCATTCCTCCAACTG 642

RESULT 12
BE744242/c 701 bp mRNA linear EST 15-SEP-2000
LOCUS 601576428F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3837367 5',
DEFINITION mRNA sequence.
ACCESSION BE744242
VERSION BE744242.1 GI:10158234
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 701)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNC522 row: n column: 08
High quality sequence stop: 666.
Location/Qualifiers

FEATURES
source 1..701
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3837367"
/issue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_9"
/notes="Organ: ovary; Vector: pOTB; Site 1: XhoI; Site 2:
EcoRI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 6.5%; Score 324; DB 2; Length 701;
Best Local Similarity 73.4%; Pred. No. 2,3e-18;
Matches 477; Conservative 0; Mismatches 125; Indels 48; Gaps 3;
Qy 343 TCTTCTTTTTTTTTTTTTTTTTTTTTTTGAGACGAGTCTGCTGCGCCAG 402
Db 688 TTTTCTTTTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTCTGCTGCTCCAG 629

Qy 403 GCTGAGTGCAGTGGCGGAGTCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCA 462
Db 628 GCTGAGTGCAGTGGCGGAGTCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCA 569
Qy 463 TTCTCTGCTCAAGCTCCGCAAGTGGAGTCAAGGAGCCGCTCAAGCCCGCT 522
Db 568 TTCTCTGCTCAAGCTCCGCAAGTGGAGTCAAGGAGCTCCGCTCAAGCCCGCT 509
Qy 523 AATT---TTTTGTATTTTATAGTGAAGAGCGGGTTTCAAGCTCCGCTCAAGCT 579
Db 508 AATTCTTTTATTTTATTTTATAGTGAAGAGCGGGTTTCAAGCTCCGCTCAAGCT 449
Qy 580 GATCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 639
Db 448 GATCTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 389
Qy 640 GAGCCACCGGCGCCCGGCTATGATCTTTC----- 669
Db 388 GAGTCACTCTGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAT 329
Qy 670 --TTGACTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 727
Db 328 GGTATGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
Qy 728 AGTTTGGGCGGAGCACCGGCTCATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 787
Db 268 CAAGTGGGCGGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209
Qy 788 GGTGAATCACTTGAAGGCGAGAGTTTGAAGCAGCTGCGCAATGCTGCTGCTGCT 847
Db 208 GGTGAATCACTTGAAGGCGAGAGTTTGAAGCAGCTGCGCAATGCTGCTGCTGCTGCT 149
Qy 848 TTTTCTTAAAAAATCAAAAAAATCAAGCGGGTTCGTGGGCACTGCTGA 907
Db 148 TTTTCTTAAAAAATCAAAAAAATCAAGCGGGTTCGTGGGCACTGCTGA 102
Qy 908 ATCCGAGCTATGCTGAGGCTGAGGCGAGAGTCACTTGAACCTGAG 957
Db 101 ATCCGAGCTATGCTGAGGCGAGAGTCACTTGAACCTGAG 52

RESULT 13
HSM807323/c 5511 bp mRNA linear HTC 22-SEP-2004
LOCUS HSM807323
DEFINITION Homo sapiens mRNA; CDNA DKFZp686N01223 (from clone DKFZp686N01223).
ACCESSION BX647179
VERSION BX647179.1 GI:34366207
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 5511)
AUTHORS Wambutt, R., Heubner, D., Mewes, H.W., Well, B., Amdt, C., Osaenger, A.,
Rodo, G., Han, M. and Wiemann, S.
CONSRM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuburg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by Agova (Berlin/Germany) within the CDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686N01223) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686N01223
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source 1..5511
/organism="Homo sapiens"
/mol_type="mRNA"

Db	5696	CACCATGCGCTGGCCCTGCTGCTACTATGTTTGTGTATGTTGCCACAGAGTGTCTGAGTATCTCAC	57555
Qy	671	TGACTATGCTGATGTGATGACAAAGTACCTAAAGCCATCAGATC-----TACCCTT	718
Db	5756	TCCCTCCCTGCTGTGGGCTCCCTTTGGTCAATTTCCCAAAATCCTATTACAGTTTACCAAT	5815
Qy	719	TAAATATGCACTTTGGGCGCAGGACCGTGGCTCATGCTGTATATTCAGCACTTTGGGAG	778
Db	5816	ATMAAAGTTGTGTGGGCGAGGCATATGGGCTCAAGCCTGTGTGATCCCGGCAATTTTGGGAG	5875
Qy	779	GCAGAGTGGGGGATTCACCTTGAGGCGCAGAGTTTGACACGACCTGGCCCAACATGCTGA	838
Db	5876	GCGAGGGCGGGTGATTCAGTAGAGGTCAAGAACTTAAACACGCTGGCCCAACATGGTGA	5935
Qy	839	AACCTGTCTTACTACPAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCGGGGTGTGCGTGGGCG	898
Db	5936	AAACCCGCTCTACT-----AAAAAAAAAAAAAAAAATTAAGTCTGGGGCTGTGGGCGAG	5984
Qy	899	ACACCTGTAAATCCAGCTATGCTGTGAGGCTGAGGACGAGATCACTTGAACCCCTGAGG	958
Db	5985	GCGCCTGTAGTCCAGCTGTGCTCGGAGGCTGAGGACGAGAAATTCTTGAAGCCCGGAGG	6044
Qy	959	CGGAGGTTGCACTGGGCGCGAGATCAATCAACCGCCTTCAAGCTTGGGCGACAGACAAAGA	1018
Db	6045	TAGAGGTTGCACTGAGCCGAGACTGCTCATCTTGCACCTCAGGCTTGGGCGACAGAGCTAGA	6104
Qy	1019	CTCTGTCTCAATAATTAATTAATTAACAAACGAA	1050
Db	6105	CACCATCTCAAAATATATTAACCAATTAAGAAG	6136

RESULT 15	
LOCUS	BC038630
DEFINITION	BC038630 Homo sapiens, similar to hypotheoretical protein FLJ20489, clone IMAGE:5263792., mRNA.
ACCESSION	BC038630
VERSION	BC038630.1 GI:24116283
KEYWORDS	HTC.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (Bases 1 to 2772) Strauberg,R. Direct Submission
AUTHORS	
TITLE	
JOURNAL	Submitted (15-Oct-2002) National Institutes of Health, Mammalian

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

FEATURES

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"
```

```

/c1one="IMAGE:5263792"
/c1ssue_type="Brain/ hippocampus"
/c1ssue_lib="NIH_MGC_95"
/lab host="DH10B"
/note="Vector: pbluescript"

```

	Query Match Similarity	6.4%; Score 321.6; DB 3; Length 2772;
	Best Local Similarity	72.4%; Pred. No. 1.5e-18;
	Matches	537; Conservative 0; Mismatches 174; Indels 31; Gaps 8;
OY	327	GNGGTGCTATGATCATCTTCTTTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTC 386
Dd	2038	GTGAATTACATAAACCCCTTTCTTTCTTTCTTTCTTTCTTTCTTTTGTGAGATGSAATC 2097
OY	387	TCCGCTCTGCGCCAAGGTGAGATGCAATGGCGGAACTCGGCTCACATGCAAGCTCCGC 446
Dd	2098	TTCTCTGTGGCCCAAGGCT-GAGTGCAATGTGTGATCTCGGCTCACATGCAACCTTAAC 2156
OY	447	TCCGGGGTTCACGCATCTCTCTCTCAGGCTCCCAAATGACTGGGACTACAGGCGCC 506
Dd	2157	TCTGTGGTTACAAGCAATCTCTCTCTCAGGCTCCGAGTAGCTGGGCTACAGGCAATC 2216
OY	507	GCCAATACGCCGGCTAATTTTTTTTGTATTTTGTAGAGACGGGGTTTACCCTTTTAA 566
Dd	2217	ACCACTACGCGCTGGCTGA--TTTTTGTATTTTGTAGAGGCAAGTTTACCAATGTTGG 2274
OY	567	CCGGGATGCGCTGCATCTCTGAACTC--GTGATCCGCGCGCTCCGCTCCCAAGTC 624
Dd	2275	TCAGGCTGG-TTTAACTCTGAACTTACAGGTACCCGCTCCGCTCTCAAGGTG 2333
OY	625	TGGATTTACAGGCGCTGAGCCACCGCGCGCGCATGATCATCTTCTGACTATGC----- 679
Dd	2334	TGTGATTTACAGGATGAGCCACCACTCTGSCATATAACATTTTCAGTAATTCATCT 2393
OY	680	-----TGATGTGACAACTAAGTAAAGCATCACTCACTTTAATATGAG--TTT 732
Dd	2394	CTATTTAATTTTGGCAGAGTTCACTTTTACACCAAGATAGTGAATATACAAAGATCC 2453
OY	733	GGGCAAGGCAACGATGCTCATGCTGTAAATTCAGACACTTTGGAGGAGAGTGGGTA 792
Dd	2454	TAGCTGGAGCTGTGGCTCATGCTGTAAATCCACCACTTTGGAGGCTGGGCGAGTGG 2513
OY	793	ATCATTTGAGGCCAGAAGTTGAGACCAAGCTGGCCAAATGATGAACTCTGTCTTAC 852
Dd	2514	ATCATCTGAGGTCAAGAGTTGAGACCAAGCTGGCCAAATGATGAAATCC----- 2564
OY	853	TAAAAAATTTTTTTTTTTTTTTTTTTTTTATCAGCCGGGTGTCTGTGGGACACTGTAAATCC 912
Dd	2565	---CAAAATATCTTAAATATACAGAAATTTGGCTGGGCTGTGGGCGGCCACTGTGTGATCC 2621
OY	913	AGCTATGCTGAGGCTGAGGCAAGAGTCACTTGAACCTTGAGAGGCGGAAGTTGCAATG 972
Dd	2622	GGCTACTTTGGAGGCTGAGGCAAGAGATCTGTTGAACCCGGAGAGTCAATTTGCAAG 2681
OY	973	GGCGGAGATCAATCACCGCCCTTCAGCTGGGCGAC-AGAGCAAGACTGTCTCAAT 1031
Dd	2682	AGCGGAGATGCAACATTTGCACTCAAGCTGGAGAGACAAAGCAAACTCGTCTTACCA 2741
OY	1032	AAATTAATTAACAAACGACCA 1053
Dd	2742	AAAAAAAAAAAAAAAAAAAAA 2763

Search completed: February 27, 2005, 12:42:13
Job time : 15686.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:13 ; Search time 310.935 Seconds
(without alignments)
10530.153 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gaaacccaagagcccccaag.....tcgagggggaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 8181359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptcdat1/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptcdat1/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptcdat1/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptcdat1/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptcdat1/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptcdat1/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1999.4	99.9	10818	4 US-09-949-016-13583	Sequence 13583, A
2	601	30.0	601	4 US-09-949-016-63402	Sequence 63402, A
3	600.6	30.0	601	4 US-09-949-016-63399	Sequence 63399, A
4	600.6	30.0	601	4 US-09-949-016-63400	Sequence 63400, A
5	600.6	30.0	601	4 US-09-949-016-63401	Sequence 63401, A
6	600.6	30.0	601	4 US-09-949-016-63403	Sequence 63403, A
7	600.6	30.0	601	4 US-09-949-016-63404	Sequence 63404, A
8	599	29.9	601	4 US-09-949-016-63405	Sequence 63405, A
9	240.2	12.0	55827	3 US-09-813-133A-3	Sequence 3, Appl1
10	240.2	12.0	55827	4 US-10-212-877-3	Sequence 3, Appl1
11	235.6	11.8	75674	4 US-09-949-016-17557	Sequence 17559, A
12	233.2	11.5	265038	4 US-09-949-016-15779	Sequence 15779, A
13	229.8	11.5	131724	4 US-09-949-016-12893	Sequence 12893, A
14	226.2	11.3	109159	4 US-09-949-016-14169	Sequence 14169, A
15	226.2	11.3	109159	4 US-09-949-016-14170	Sequence 14170, A
16	225.6	11.3	18319	4 US-09-949-016-17446	Sequence 17446, A
17	225.4	11.3	325791	4 US-09-768-185A-1	Sequence 1, Appl1
18	225.2	11.3	601	4 US-09-949-016-202742	Sequence 202742, A
19	224.6	11.2	72992	4 US-09-949-016-17552	Sequence 17592, A
20	221.4	11.1	70131	4 US-09-949-016-16064	Sequence 16064, A
21	221.2	11.1	150394	4 US-09-949-016-13042	Sequence 13042, A
22	220.8	11.0	209210	4 US-09-949-016-15094	Sequence 15094, A
23	220	11.0	136264	4 US-09-949-016-12756	Sequence 12756, A
24	220	11.0	136265	4 US-09-949-016-13001	Sequence 13001, A
25	219.8	11.0	166698	4 US-09-949-016-16038	Sequence 16038, A
26	219.4	11.0	35058	4 US-09-949-016-12607	Sequence 12607, A
27	219.4	11.0	35059	4 US-09-949-016-13831	Sequence 13831, A

28	219.4	11.0	165651	4 US-09-949-016-13032	Sequence 13032, A
29	219.2	11.0	169684	4 US-09-949-016-16893	Sequence 16893, A
30	218.6	10.9	17353	4 US-09-949-016-13437	Sequence 13437, A
31	218.6	10.9	301828	4 US-09-949-016-13369	Sequence 13369, A
32	218.4	10.9	15358	4 US-09-949-016-13382	Sequence 13382, A
33	218.2	10.9	91831	4 US-09-949-016-13694	Sequence 13694, A
34	218	10.9	88490	4 US-09-949-016-12758	Sequence 12758, A
35	218	10.9	88766	4 US-09-949-016-14232	Sequence 14222, A
36	217.8	10.9	601	4 US-09-949-016-158456	Sequence 58456, A
37	217.8	10.9	30371	4 US-09-949-016-15385	Sequence 15395, A
38	217.4	10.9	44789	4 US-09-949-016-15309	Sequence 13909, A
39	217.2	10.9	47751	4 US-09-949-016-12812	Sequence 12812, A
40	217.2	10.9	47752	4 US-09-949-016-16955	Sequence 16955, A
41	216.8	10.8	107421	4 US-09-949-016-15532	Sequence 15532, A
42	216.4	10.8	99787	4 US-09-949-016-15255	Sequence 15255, A
43	216.4	10.8	256171	4 US-09-949-016-12822	Sequence 12822, A
44	216.4	10.8	256176	4 US-09-949-016-15524	Sequence 15524, A
45	216.2	10.8	8230	4 US-09-949-016-14183	Sequence 14183, A

ALIGNMENTS

RESULT 1					
US-09-949-016-13583					
Sequence 13583, Application US/09949016					
Patent No. 6812339					
GENERAL INFORMATION:					
APPLICANT: VENTER, J. Craig et al.					
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED					
FILE REFERENCE: C1001307					
CURRENT APPLICATION NUMBER: US/09/949,016					
CURRENT FILING DATE: 2000-04-14					
PRIOR APPLICATION NUMBER: 60/241,755					
PRIOR FILING DATE: 2000-10-20					
PRIOR APPLICATION NUMBER: 60/237,768					
PRIOR FILING DATE: 2000-10-03					
PRIOR APPLICATION NUMBER: 60/231,498					
PRIOR FILING DATE: 2000-09-08					
NUMBER OF SEQ ID NOS: 207012					
SOFTWARE: FASTSEQ for Windows Version 4.0					
SEQ ID NO 13583					
LENGTH: 10818					
TYPE: DNA					
ORGANISM: Human					
US-09-949-016-13583					
Query Match					
Best Local Similarity 99.9%; Score 1999.4; DB 4; Length 10818;					
Matches 2000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					
QY	1	GACACCAAGAGCCCCCAAGAGTGTCCAGGTTCTTGTGATACGACCACTCTGTC	60		
DB	3245	GACACCAAGAGCCCCCAAGAGTGTCCAGGTTCTTGTGATACGACCACTCTGTC	3304		
QY	61	CGCCTTCACATCTTCTTCACATTTCTCATTTGTTCTCTTTGAAGTGCATCTCCTAAG	120		
DB	3305	CGCCTTCACATCTTCTTCACATTTCTCATTTGTTCTCTTTGAAGTGCATCTCCTAAG	3364		
QY	121	TCTCCAGGCTTGAAATATCAGGCTGTAAGTCAGAAAAGTGCAGAGATCTTACCCGA	180		
DB	3365	TCTCCAGGCTTGAAATATCAGGCTGTAAGTCAGAAAAGTGCAGAGATCTTACCCGA	3424		
QY	181	GACAGATAGATGACACCATTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGTCAGG	240		
DB	3425	GACAGATAGATGACACCATTTCTGCGCGGTGACAAAGAGGTAGAGACTCTGTCAGG	3484		
QY	241	TGAGACACTCTCTTTTATTCAGACATACACTGATGCCAATCTCGGTAATGAGAGC	300		
DB	3485	TGAGACACTCTCTTTTATTCAGACATACACTGATGCCAATCTCGGTAATGAGAGC	3544		
QY	301	GTTGCCAAATTCTGAGATTCAGCAATTCGCAAGACAGTCCAGACCCCTGTTCTCAGAG	360		

3545 GTTGCAAAATTTGAGAAATCCAAATTCGCAAGACGTCAGGACCCCTGTTCTCACAGA 3604
361 GCTGATACCTTAAGTAGAGTGTGTTTAGTAGAAATAATGCTGAGCTGCTTAATGCTATTC 420
3605 GCTGATACCTTAAGTAGAGTGTGTTTAGTAGAAATAATGCTGAGCTGCTTAATGCTATTC 3664
421 CAGTTTTTTAGTAGCAATTTAAACAGGTAATAAGGCTGGCGCAGTGGCTCACACCT 480
3665 CAGTTTTTTAGTAGCAATTTAAACAGGTAATAAGGCTGGCGCAGTGGCTCACACCT 3724
481 GTAATCCAGACCTTTGGAGGCTGAGGCAAGCAGATCACCTTTGGTCAGAGCTTTGAGA 540
3725 GTAATCCAGACCTTTGGAGGCTGAGGCAAGCAGATCACCTTTGGTCAGAGCTTTGAGA 3784
541 CTGAGCTGGCCAAATGCGGAAATCTGCTCTTAAATAATAATGCAAAATTTGAGCTGGC 600
3785 CTGAGCTGGCCAAATGCGGAAATCTGCTCTTAAATAATAATGCAAAATTTGAGCTGGC 3844
601 ATGATGGCGGCGCTGTATCTCAGCTGCTCAGAGCCGAGACACAAGAAATCACTTAA 660
3845 ATGATGGCGGCGCTGTATCTCAGCTGCTCAGAGCCGAGACACAAGAAATCACTTAA 3904
661 ACCAAGAGGTGAGGTTGCACTGAGCTGAGCTGATCTGCTCACTTCCAACTGGAGAC 720
3905 ACCAAGAGGTGAGGTTGCACTGAGCTGAGCTGATCTGCTCACTTCCAACTGGAGAC 3964
721 AGAGTACACCTTTGCTCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 780
3965 AGAGTACACCTTTGCTCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 4024
781 TTAACTTAAATTAACCAATGTATCCAAATATCAATCTTCAAGTGTATTAATTAATA 840
4025 TTAACTTAAATTAACCAATGTATCCAAATATCAATCTTCAAGTGTATTAATTAATA 4084
841 ACAATTAATGATGATGATCTTAAATCTTCTTTCTGTTTCAATTAATGCTTTGAAAGT 900
4085 ACAATTAATGATGATGATCTTAAATCTTCTTTCTGTTTCAATTAATGCTTTGAAAGT 4144
901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGGAATGATCTTCAATTTGAGGTC 960
4145 GAGTATATATGTTATGCTGACAGACATCTCAATTTGGAATGATCTTCAATTTGAGGTC 4204
961 AGTAGCACATGTTGCTGAGCAGTGTACTGTATTTGATGAGCAGGATCTAGAGGAAAGATC 1020
4205 AGTAGCACATGTTGCTGAGCAGTGTACTGTATTTGATGAGCAGGATCTAGAGGAAAGATC 4264
1021 AGGGCTGTTTGTATGTTGGGCAAGTTGTCAGCTGATTAAGAAATCAATATCTTAATAG 1080
4265 AGGGCTGTTTGTATGTTGGGCAAGTTGTCAGCTGATTAAGAAATCAATATCTTAATAG 4324
1081 GGCACCTCCGTTACAGATGTCAGTTTGGCACTTTTTCAGGCGTGTGTAAAGTTC 1140
4325 GGCACCTCCGTTACAGATGTCAGTTTGGCACTTTTTCAGGCGTGTGTAAAGTTC 4384
1141 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTCTGGTAAATATCTTGA 1200
4385 TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTCTGGTAAATATCTTGA 4444
1201 GGAAGAAAGAGAAATCTGTATGTTTTCAGAGAAATTTTAAATACAGGGAATTA 1260
4445 GGAAGAAAGAGAAATCTGTATGTTTTCAGAGAAATTTTAAATACAGGGAATTA 4504
1261 ATTGCAAGGCTGCGAGAGGCTGAGGAAACAAGTTAAATAATAATAATACTGTGTC 1320
4505 ATTGCAAGGCTGCGAGAGGCTGAGGAAACAAGTTAAATAATAATAATACTGTGTC 4564
1321 AAGATCTGATTAATAGGCAATTTCAAGAGTGTGAAGTTAAATCCCAATATAAAC 1380
4565 AAGATCTGATTAATAGGCAATTTCAAGAGTGTGAAGTTAAATCCCAATATAAAC 4624
QY ATGTTTTTAGATATGAATCAATTAAGGCAATATTTCAAAAGGTGTCAGGGAGCTTC 1440

4625 ATGTTTTTAGATATGAATCAATTAAGGCAATATTTCAAAAGGTGTCAGGGAGCTTC 4684
QY 1441 CTGAGAGAGTGGCAATTTGAGAGAGAAATGATGACACAAAGAGCTTAATCTGTAGAT 1500
DB 4685 CTGAGAGAGTGGCAATTTGAGAGAGAAATGATGATGACACAAAGAGCTTAATCTGTAGAT 4744
QY 1501 TTAAGGGGAAAGAAAGGACGCTGCAAGGCTGAGGCAATTAAGTGTGCTGATTC 1560
DB 4745 TTAAGGGGAAAGAAAGGACGCTGCAAGGCTGAGGCAATTAAGTGTGCTGATTC 4804
QY 1561 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAGTGGGGCAACATGTAAGATG 1620
DB 4805 AAAAGAAAGAGAAACCAATGCACTGAGAAACAAAGTGGGGCAACATGTAAGATG 4864
QY 1621 AGGCTGAGAGTGTAGGCAAGGCGAATGCTGCAAGTATTTCTTGGTCAACACAGA 1680
DB 4865 AGGCTGAGAGTGTAGGCAAGGCGAATGCTGCAAGTATTTCTTGGTCAACACAGA 4924
QY 1681 GCTTCCCTATGTTCTTAATGAGCTGTATCTGTTGAGAGACAGAAATTTAAATCAAC 1740
DB 4925 GCTTCCCTATGTTCTTAATGAGAGCTGTATCTGTTGAGAGACAGAAATTTAAATCAAC 4984
QY 1741 TGTATCATTAACACAGACCTTCTCTGTATTTAGAGCTCCCAAGGATCTTAAGAGCTTA 1800
DB 4985 TGTATCATTAACACAGACCTTCTCTGTATTTAGAGCTCCCAAGGATCTTAAGAGCTTA 5044
QY 1801 AGTTAACAAAGCTCTCATTTAGCAGGAGTGTGTTTCAACAGTATTAAGAAAGTGGAGAT 1860
DB 5045 AGTTAACAAAGCTCTCATTTAGCAGGAGTGTGTTTCAACAGTATTAAGAAAGTGGAGAT 5104
QY 1861 CAGAGTACTCCAGTCCCAATGCTATGAAAGCTCCCAATTTGTAACAACTTGACAA 1920
DB 5105 CAGAGTACTCCAGTCCCAATGCTATGAAAGCTCCCAATTTGTAACAACTTGACAA 5164
QY 1921 ATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGTCCTGAGTGTGGGGGGTGG 1980
DB 5165 ATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGTCCTGAGTGTGGGGGGTGG 5224
QY 1981 TTGCGAGGGGAAACTTTTA 2001
DB 5225 TTGCGAGGGGAAACTTTTA 5245

RESULT 2
US-09-949-016-63402/C
; Sequence 63402, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 63402
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63402

Query Match 30.0%; Score 601; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.7e-147;
Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 456 AGGCTGGCGCAGTGGCTCAGACCTGTATCCAGCAGCTTTGGAGGCTGAGGCAAGCAG 515

Db 601 AGGCTGGGCGGAGTGGCTCAACCTGTATCCAGCACTTTGGAGGTGAGGAGGAG 542
Qy 516 ATCACTTTGGTCAAGAGTTTGAAGCTAGTACCTGGCCAACTGGGAACTCTGTCTTAA 575
Db 541 ATCACTTTGGTCAAGAGTTTGAAGCTAGTACCTGGCCAACTGGGAACTCTGTCTTAA 482
Qy 576 AAAAAATPACAAAATTAAGCTGGATGAGTGGCGGGGCGCTGTATCTCAAGCTGCTAG 635
Db 481 AAAAAATPACAAAATTAAGCTGGATGAGTGGCGGGGCGCTGTATCTCAAGCTGCTAG 422
Qy 636 AGGCGGAGACACAAAGATCACTTAAACCCAGAGAGTGGAGGTGCTAGTGAAGTGC 695
Db 421 AGGCGGAGACACAAAGATCACTTAAACCCAGAGAGTGGAGGTGCTAGTGAAGTGC 362
Qy 696 TGGCACTCACTCCAACTGGGAGACAGAGTGACATTTTGTCTCAAAAAAGAAAAAAA 755
Db 361 TGGCACTCACTCCAACTGGGAGACAGAGTGACATTTTGTCTCAAAAAAGAAAAAAA 302
Qy 756 ACAAGTAAAAAGAAACAGGTGAAGTTAACTTTAATTAACCAATGTATCCAAATACAA 815
Db 301 ACAAGTAAAAAGAAACAGGTGAAGTTAACTTTAATTAACCAATGTATCCAAATACAA 242
Qy 816 CATTTCAAGGTATTAATTAATTAACCAATTAATGAATGAGTACTTAACTTTCTT 875
Db 241 CATTTCAAGGTATTAATTAATTAACCAATTAATGAATGAGTACTTAACTTTCTT 182
Qy 876 GTTTTCATTAATTAAGTCTTTGAAAGTGAATATATATGTATGCTGACAGCATCTCAAT 935
Db 181 GTTTTCATTAATTAAGTCTTTGAAAGTGAATATATATGTATGCTGACAGCATCTCAAT 122
Qy 936 TGdACTGATCACTTACAGGCTCAGTACGCAATGTGGTACACTTACTGATTTGA 995
Db 121 TGdACTGATCACTTACAGGCTCAGTACGCAATGTGGTACACTTACTGATTTGA 62
Qy 996 TGdACGAGATCTAGAGGAAAGATCAGGCTGTTTGTATGTTGGCAGGTTGCACT 1055
Db 61 TGdACGAGATCTAGAGGAAAGATCAGGCTGTTTGTATGTTGGCAGGTTGCACT 2
Qy 1056 G 1056
Db 1 G 1

RESULT 3
US-09-949-016-63399/c
; Sequence 63399, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14, 755
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63399
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63399

Query Match 30.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 148 TAAGTCAGAAAAGGTGAGAGTGTCTTACCCGAGACAGATAGATGACACCATGTTCTGC 207
Db 601 TAAGTCAGAAAAGGTGAGAGTGTCTTACCCGAGACAGATAGATGACACCATGTTCTGC 542
Qy 208 CCGGTGACAAAGAGGTAGAGTCTCTGCGAGGTGAGACACCTCTTTTATTCAGCAGA 267
Db 541 CCGGTGACAAAGAGGTAGAGTCTCTGCGAGGTGAGACACCTCTTTTATTCAGCAGA 482
Qy 268 TACACACTGATGCTCACTGTTAATGATGAGAGGTGCTCCAAATTTGAGAAATCCAGAA 327
Db 481 TACACACTGATGCTCACTGTTAATGATGAGAGGTGCTCCAAATTTGAGAAATCCAGAA 422
Qy 328 TGGCAAGACAGTACAGGACCCCTGTTCTACAGAGCTCATTCCTAGAGTATGAGTTTA 387
Db 421 TGGCAAGACAGTACAGGACCCCTGTTCTACAGAGCTCATTCCTAGAGTATGAGTTTA 362
Qy 388 GTAGAAATTAATGCTGAGCTGCTTATGATTTCACTTTTCACTTTTATAGTACCAATTAAACA 447
Db 361 GTAGAAATTAATGCTGAGCTGCTTATGATTTCACTTTTCACTTTTATAGTACCAATTAAACA 302
Qy 448 GGTAAAAAGGCTGGGCGAGTGGCTCACTCTGTAATCCAGCACTTTGGAGGCTGAG 507
Db 301 GGTAAAAAGGCTGGGCGAGTGGCTCACTCTGTAATCCAGCACTTTGGAGGCTGAG 242
Qy 508 GCAAGGAGATCACTTTGCTGAGAGTTGAGACTGAGTGGCTGAGCAATGCGGAACTCT 567
Db 241 GCAAGGAGATCACTTTGCTGAGAGTTGAGACTGAGTGGCTGAGCAATGCGGAACTCT 182
Qy 568 GTCTTAAAAAAAATACAAAAATTAGCTGAGTGGCTGGGCGCTGTTAATCTCAGC 627
Db 181 GTCTTAAAAAAAATACAAAAATTAGCTGAGTGGCTGGGCGCTGTTAATCTCAGC 122
Qy 628 TGCTCAGAGGCGGAGACACAAATCACTTAAACCCAGAGGTGAGAGTTGAGTGAAC 687
Db 121 TGCTCAGAGGCGGAGACACAAATCACTTAAACCCAGAGGTGAGAGTTGAGTGAAC 62
Qy 688 TAGATGATGCTCACTCACTGAGAGTGGAGACAGAGTGACATTTTGTCTCAAAAAAGAA 747
Db 61 TAGATGATGCTCACTCACTGAGAGTGGAGACAGAGTGACATTTTGTCTCAAAAAAGAA 2
Qy 748 A 748
Db 1 A 1

RESULT 4
US-09-949-016-63400/c
; Sequence 63400, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14, 755
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63400
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-63400

Query Match 30.0%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 2.1e-147;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

1      RESULT 5
2      US-09-949-016-63401/c
3      ; Sequence 63401, Application US/09949016
4      ; Patent No. 6812339
5      ; GENERAL INFORMATION:
6      ; APPLICANT: VENTER, J. Craig et al.
7      ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
8      ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
9      ; FILE REFERENCE: C1001307
10     ; CURRENT APPLICATION NUMBER: US/09/949, 016
11     ; CURRENT FILING DATE: 2000-04-14
12     ; PRIOR APPLICATION NUMBER: 60/241,755
13     ; PRIOR FILING DATE: 2000-10-20
14     ; PRIOR APPLICATION NUMBER: 60/237,768
15     ; PRIOR FILING DATE: 2000-10-03
16     ; PRIOR APPLICATION NUMBER: 60/231,498
17     ; PRIOR FILING DATE: 2000-09-08
18     ; NUMBER OF SEQ ID NOS: 207012
19     ; SOFTWARE: FastSeq for Windows Version 4.0
20     ; SEQ ID NO 63401
21     ; LENGTH: 601
22     ; TYPE: DNA
23     ; ORGANISM: Human
24     ; US-09-949-016-63401

```

Matches	600;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	242	GAGGACACCTCTCTTTATATTCAGCAGATACACACTGAGTGGCACTCGTGTAAACAATGAGGG	301						
Db	601	GAGGACACCTCTCTTTATATTCAGCAGATACACACTGAGTGGCACTCGTGTAAACAATGAGGG	542						
QY	302	TTGGCCAAATTCCTGAGAAATCCAGCAATTCGCAAGACAGTCAAGACCCCTGTCTCAACAGAG	361						
Db	541	TTGGCCAAATTCCTGAGAAATTCAGCAATTCGCAAGACAGTCAAGACCCCTGTCTCAACAGAG	482						
QY	362	CTCATACCTTGAAGTAGTGTGTTTAAATATAATATGCTGAGCTGCTTAATGTCAATTTCC	421						
Db	481	CTCATACCTTGAAGTAGTGTGTTTAAATATAATATGCTGAGCTGCTTAATGTCAATTTCC	422						
QY	422	AGTTTTTTAATGATGACCACTTAAAAACAGTAAAAAAAGCTGGGGCAGTGGCTCAACCTGG	481						
Db	421	AGTTTTTTAATGATGACCACTTAAAAACAGTAAAAAAAGCTGGGGCAGTGGCTCAACCTGG	362						
QY	482	TAAATCCAGCACTTTGGGAGGCTGAGGCGAGGACAGATCACTTGTGTACAGAGTTTGAGAC	541						
Db	361	TAAATCCAGCACTTTGGGAGGCTGAGGCGAGGACAGATCACTTGTGTACAGAGTTTGAGAC	302						
QY	542	TAGCCTGGCCAAACATGCGGAATCTGTCTCTAAAAAAAATACAAAATTAAGCTTGACA	601						
Db	301	YAGCCTGGCCAAACATGCGGAATCTGTCTCTAAAAAAAATACAAAATTAAGCTTGACA	242						
QY	602	TGTGTGGCGGGCGCTGTATATCTCAGCTGCTCAGAGGCGGACACACAGATACCTTAAA	661						
Db	241	TGTGTGGCGGGCGCTGTATATCTCAGCTGCTCAGAGGCGGACACACAGATACCTTAAA	182						
QY	662	CCAGAGAGTGGAGGTTGSCAGTSGAGTGCAGATGCTGTCACCTCACTCAACTGGAGACA	721						
Db	181	CCAGAGAGTGGAGGTTGSCAGTSGAGTGCAGATGCTGTCACCTCACTCAACTGGAGACA	122						
QY	722	GAGTGACACTTTTGTCTCAAAAAGAAAAAAAACAGTAAAAAAAGAAACAGGTGAAGT	781						
Db	121	GAGTGACACTTTTGTCTCAAAAAGAAAAAAAACAGTAAAAAAAGAAACAGGTGAAGT	62						
QY	782	TAACTTTAATAACCCAAATGATATCCCAATATCAATCATTTCAAAAGTGAATTAATATAAA	841						
Db	61	TAACTTTAATAACCCAAATGATATCCCAATATCAATCATTTCAAAAGTGAATTAATATAAA	2						
QY	842	C 842							
Db	1	C 1							

RESULT 6
 US-09-949-016-63403/C
 ; Sequence 63403, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CLO01307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 63403
 ; LENGTH: 601
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-63403

30.0%, Score 600.6; DB 4; Length 601;

Query Match 29.9%; Score 599; DB 4; Length 601;
Best Local Similarity 99.7%; Pred. No. 5.6e-147;
Matches 599; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1301 AAAAAAAAACTGTGTGTCAAGATCTGCATTAATAGGCAATTCAGAGAGTGTAA 1360
DB AAAAAAAAACTGTGTGTCAAGATCTGCATTAATAGGCAATTCAGAGAGTGTAA 542
QY 1361 GGTAAACCCCAATTAATATGTTTTAGATATGTAACATTAAGGCCAATTTCAA 1420
DB 541 GGTAAACCCCAATTAATATGTTTTAGATATGTAACATTAAGGCCAATTTCAA 482
QY 1421 AAGGTGTCAAGGAGCTCTTGGAGAGGTGATTTGAGCAAGAAATGATGACAA 1480
DB 481 AAGGTGTCAAGGAGCTCTTGGAGAGGTGATTTGAGCAAGAAATGATGACAA 422
QY 1481 AGAAGCTAACTGTGAAGTTTAAAGGGAAAGAAAGGCACTGCAAAAGCCCTGAGCA 1540
DB 421 AGAAGCTAACTGTGAAGTTTAAAGGGAAAGAAAGGCACTGCAAAAGCCCTGAGCA 362
QY 1541 GTAGGAATTTGGCTGATTTCAAGAAAGAAAGCAATGCACTGGAACAAAGT 1600
DB 361 GTAGGAATTTGGCTGATTTCAAGAAAGAAAGCAATGCACTGGAACAAAGT 302
QY 1601 GGGGCAACAGTAAAGTGAAGTGTAGGAGGGGCAATGCTGTGCAAGTAT 1660
DB 301 GGGGCAACAGTAAAGTGAAGTGTAGGAGGGGCAATGCTGTGCAAGTAT 242
QY 1661 TTCTTGTCACCAACACAGAGCTTCCCTATGTTCTAATGAGCTGATCTGTAGGAA 1720
DB 241 TTCTTGTCACCAACACAGAGCTTCCCTATGTTCTAATGAGCTGATCTGTAGGAA 182
QY 1721 GAAGAATTTAAATCAAACTGTATCAATCAACAGACCTTCTCTGTATTCAGGCTCCC 1780
DB 181 GAAGAATTTAAATCAAACTGTATCAATCAACAGACCTTCTCTGTATTCAGGCTCCC 122
QY 1781 AAGGATCTGAAGAGAGTAACTTAACAAGCTCTCAATTAAGAGGAGTGTTCACAG 1840
DB 121 AAGGATCTGAAGAGAGTAACTTAACAAGCTCTCAATTAAGAGGAGTGTTCACAG 62
QY 1841 TAGTTAGAAAGTGGGAGTTCAGAGTACTCCAGTCCCATGCTATGAAAAAGTCCCCC 1900
DB 61 TAGTTAGAAAGTGGGAGTTCAGAGTACTCCAGTCCCATGCTATGAAAAAGTCCCCC 2
QY 1901 A 1901
DB 1 A 1

RESULT 9
US-09-813-133A-3/c
; Sequence 3, Application US/09813133A
; Patent No. 6455294
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001173
; CURRENT APPLICATION NUMBER: US/09/813,133A
; PRIOR FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Human
US-09-813-133A-3

Query Match 12.0%; Score 240.2; DB 3; Length 55827;
Best Local Similarity 68.6%; Pred. No. 4e-52;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTAGAGTAGTGGTTTGTAGAGAAATATCTGAGCTGCTTATGTCTTCCA 422
DB 46943 TGAATCTCAGACAGAGGTGTGTTAATAGAAATATATGACACTCATATTTACTTTAA 46884
QY 423 GTTTTATAGACCATTAATAACAGGTAAAT--AAGGTGGCGCAGTGGCTCACACCT 480
DB 46883 ATTTTCAATTTACACTTTTAAAGAAATATTAAGAGGCAAGGCAAGTGGCTCACACCT 46824
QY 481 GTATATCCAGCACTTTGGAGGCTGAGGCAAGCAATCACTTTGGTCAAGAGTTTGAGA 540
DB 46823 ATATATCCAGCACTTTGGAGGCTGAGGCAAGGCGGATCACTTGAGATCAAGAGTTTGAGA 46764
QY 541 CTAGCTGGCCAAACATGGGAAACTGTCTGTAAAAAAAATATCAAAATATAGCTGGC 600
DB 46763 CAGCAGAGGCAACATGGTAAACCTCGTCTTACCAAAA--TACAAATATGCAAGG 46706
QY 601 ATGTGTGGCGCGCTGTATCTCAGCTGCTCAGAGGCGCAGACAAAGATCACTTAA 660
DB 46705 ATATGTGTGCAAGCTGTATCTCAGCTGCTCAGAGGCGTGAAGGCAAGAACTTGCTGA 46646
QY 661 ACCAGAGAGTGAAGTGTGCACTGAGCTGAGATCTGCTCACT--CACTCCAACTGGGAAG 719
DB 46645 ACCAGAGAGGCAAGAGTGTGCACTGAGCTGAGATCACTGCACTCCAGCTGTGATGA 46586
QY 720 CAGAGTGACA--CTTTGTCTCAAAAAAGAAAAAACAAGTAAAGAAACAGGTGA 778
DB 46585 CAGACAAACCTCTGCCACACACACACAAAAAATATGTAGATTAATTTCAATTA 46526
QY 779 AGTTAATTTAAATCAACCAATGATCCCAATATCAATCAATTTCAAGTAAATTAATATA 838
DB 46525 TGTATATTTAATTAACCAATATGTACAAAGATCA--TTTCAAGTAAATTAATATA 46470
QY 839 AAACAATTTAATGATGATTAATCTTTTCTTTTCTTTTCAATTAAGTCTTGA 898
DB 46469 AA--AATTAATTTCTGAATATTTTAAATTTTAACTTT--TTTCAATTAAGTCTTGA 46417
QY 899 GTAGATATATATGTTATGCTGACAGACATCTCAATTTGAGCTAGTCAATTTCAAGTGC 958
DB 46416 TCTGATGTTAATTTAACTGATGATGTCTCAATTTCAAGACAGTCAATTTTAAGCAC 46357
QY 959 TCAGTAGCCACATGTGGCTAGCAATTAATGATGATGGA 995
DB 46356 TGAATAGCCACATGTGGTTAATGAGCTAGTATGCTGGA 46320

RESULT 10
US-10-212-877-3/c
; Sequence 3, Application US/10212877
; Patent No. 6818429
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL001173DIY
; CURRENT APPLICATION NUMBER: US/10/212,877
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/813,133
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-877-3

Query Match 12.0%; Score 240.2; DB 4; Length 55827;
Best Local Similarity 68.6%; Pred. No. 4e-52;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTAGAGTAGTGGTTTGTAGAGAAATATCTGAGCTGCTTATGTCTTCCA 422

Db 46943 TGAATCTTCAGACACGAGGTTGTTTAAATAGAAATATTAATGCAATGCATATATTACTTTAA 46884
 Oy 423 GTTTTAAAGTACCCACATTAAACAGGTAAA--AAGCTGGGCGAGTGGCTCACACCT 480
 Db 46883 ATTTTCAAGTTACACTTTTAAGAAAATATAAAAAGGCGACGAGCTGCACACCT 46824
 Oy 481 GTAAATCCCGACATTTTGGAGCTGAGGCGAGATCATCTTTGGTCAGAGATTGAGA 540
 Db 46823 ATAAATCCAGCACTTTGGAGGCTGAGGAGGGGAGATCATCTTAGATCAGAGATTGACA 46764
 Oy 541 CTAGCTGGCCAAACATGGCGAAACCTGTCTCTAAAAAAAAAATACAAAATTAATGCTGGC 600
 Db 46763 CCAAGGAGGCCCAACATGTGTAAACCCCTCTTACCAAAAA--TACAAAATATGCCAGGC 46706
 Oy 601 ATGTGGCGGAGCGCCTGTATCTCAGCTGCTCAGAGGCGGAGACAACAATCACTTAA 660
 Db 46705 ATAGTGTGTACAGCCTGTATCTCCAGCTACTTGGAGGCTGAGGACGAAACTTGGCTTGA 46646
 Oy 661 ACCGAGAGGTGGAGGTTTGCATGTAGCTGATGTCGCCACT--CATCTCCAACTTGGAGA 719
 Db 46645 ACCGAGAGGCGAGAGGTTGCAGTACGTAGATCAATGACATGCACTCCAGCTTGAATGA 46586
 Oy 720 CAGAGTGACA--CTTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
 Db 46585 CAGAACAGACCCCTGCCACACACACACACACAAAAAATATGTGATTAATTCAATA 46526
 Oy 779 AGTTAACTTAAATPAAACCAATGTATCCAAATACAAATCATTTCAAAAGTAAATTAATATA 838
 Db 46525 TTGTATTTATTTAAACCAATATGTACAAAGTATCA----TTCAAGTGAACCAATATA 46470
 Oy 839 AAACAATATGAATAGATACTTAATCTTTCTTGTGTTCATATTAATGCTTTGAAA 898
 Db 46469 AA--AATTAATCTTGAATATTTAATTAAGTACTT-----TTTATATCTAAGCCTTCAAAT 46417
 Oy 899 GTGAGTATATATGTTATGTCTGACAGACACATCTCAATTTGGACTAGTACATTTCAAGTGC 958
 Db 46416 TCTGTTGTGTATTTAACATCACTCAGTGTATGTCTCAATTCAAGACAGCATTAATTGAAGCAC 46357
 Oy 959 TCAGTAGCCACATGTGGCTAGCAGTTCATGTAATTGGA 995
 Db 46356 TGGATAGCCACATGTGGTTAGTGGCTACTATGCTGGA 46320

RESULT 11
 US-09-949-016-17597/c
 ; Sequence 17597, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949, 016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 17597
 ; LENGTH: 75674
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(75674)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17597

```

Query Match 11.8%; Score 235.6; DB 4; Length 75674;
Best Local Similarity 67.4%; Pred. No. 7.3e-51;
Matches 407; Conservative 0; Mismatches 189; Indels 8; Gaps 5;

QY 362 CTCATACCCCTAGAGTAGTGAGTGTATAGTAGAAT--AATGCTAGAGCTGTATGTCAATT 419
DB 64429 CTCAGAGCTAGAGTAGTGTGTGTGTCAAGAGAAATACATGTAAAGCACAATATTTATCT 64370
QY 420 CCAAGTTTTTATAGTAGCCACATTTAAACAGTAAACAAAGGCTGGGCGCACTGCTCACACC 479
DB 64369 TAAATTTTCTAGTGTCTCTCATTTAAACAAAGTGTAAAGAGACCGGGCGCACTGTCTCACATC 64310
QY 480 TGTATATCCAGACATTTGGAGAGGCTGAGGCAAGCATATCACTTTTGTGTAGAGATTGAG 539
DB 64309 TGTATATCCAGACATTTGGAGAGGCGGAGAGCGCGAGATCAC--GAGTCAAGAGATCAAG 64252
QY 540 ACTAGCCTGGCCACATGAGCGGAACTGTGTCTTAACAAATACAAATATAGCTGG 599
DB 64251 ACCATCTGGCTTAACATGTGTAAACCCGCTCTTTACTTAAATACAAATATAGCCAG 64192
QY 600 CATGTGGCGGGCGGCTGTAACTTCAGCTGCTCAGAGGCGGAGACACAAGATCACTTA 659
DB 64191 CGGTGTGCGAGGCACTGTAGTCCAGCTGCTCGGAGGCTGGGCGAGAGATCACTGG 64132
QY 660 AACCCAGAGGTGAGGTTGTCAGTGAAGCTGAGATGTGCACT--CACTCAACCTGGAG 718
DB 64131 AACCCGAGAGGAGAGGTTGTCAGTGAACGAGATGCGGCACTGCACTTACCTGGAGT 64072
QY 719 ACAGAGTACACTTT--TGTCCAAAAGAAAAACAAAGTAAACAAAGTAAACAAAGAGTG 777
DB 64071 ACAGAGAGACTACATCTCAACAAACAAACAAACAAATTAATTAATTAAGAG 64012
QY 778 AAGTTAATTAAATTAACCAATGTATCCCAATATCAATCAATTTCAAAGTAAATATAT 837
DB 64011 TAGGTAGAAATGATTTTATAGTTTAT--CTAACCAATATATCCAAATATTTCAATTCA 63954
QY 838 AAAACAATTATGATGAGATCTTAACTTTCTTTCTTTTATATTAAGCTTTTGA 897
DB 63953 ACATGTATCAATATTAACAAATTAAGATCACTTAATTTCTTTTCACTAAGATTTTGA 63894
QY 898 AGTAGTATATATGTTATGCTGACGACATCTCAATTTGGACGTAGTACATTTGAGTG 957
DB 63893 AATTGTATATATTTTAACTTATAGCATATCTCAATTCAGAAATGCTCATAGTGTTG 63834
QY 958 CTCA 961
DB 63833 GGCA 63830

RESULT 12
US-09-949-016-15779/c
; Sequence 15779, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15779
; LENGTH: 265038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
;

```


Db 107346 CATGGTGGGGCAACCTCTGTAATCCAGCTACTCAATAGTAGGCAACAAGATCACTTG 107405
QY AACCAAGAGGTGGAGGTTGCAGTGAGCTGAGATCGTGCACCT - CACTCCAACTTGGAG 718
Db 107406 AACCCGGGAAAGGGAGGTTGCAGTGGCCGAGATCAGCCACTGCTACTCCAGCCTGTGTG 107465S
QY ACAGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAAGTGA 778
Db 107466 ACAGAGCAAGACTCTGTCTCAAAAAAAGAAAAAATCAAGAGAAAGAAACAATGGA 107525S
QY AGTTAACTTAAATTAACCCATGATGCCAAATCAATCAATTTCAAGATGAATTAATATA 838
Db 107526 AATTATTTTACTAACAATTTAATTAAACCGATGATCCAAAAGATATATCAATTTCAAAATG 107585S
QY AAACATTTATGAATGAGATCACTTTACATTCCTTTCTGTTTCATATTAAGTCTTTGAAA 898
Db 107586 CAATTCATATATATATATTAATGAGATTAATTACATCTCTCTTATATCTATGTCCTTGACA 107645S
QY GTGAGATATATATGTTATGCTGACAGACATCTCAATTTTGGAC 940
Db 107646 TCATGTGTGTGTTTACACTTACAGCCCAATTTTAATTCGAC 107687

```

RESULT 15
US-09-949-016-14170
; Sequence 14170, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14170
; LENGTH: 109159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14170

```

Query Match	11.3%	Score 226.2;	DB 4;	Length 109159;
Best Local Similarity	68.4%;	Pred. No. 2.5e-48;		
Matches 357;	Conservative	0;	Mismatches 158;	Indels 7;
				Gaps 3;

QY	435	TTTTAGTGGCCACATTTAAACAGGTA-----AAAAGCTGGGCCAGTGGTCAACC	479
Db	107167	TTTCTAGTGGCCACATTTTAAACAATTAAGAAAGAAAGGCCAGGCTGTGGTCTATGCC	107226
QY	480	TGTAATCCCAACACTTTGGGAGCGCTGAGCGAGCAGATCACTTTGGTCAAGAGTTTGAG	539
Db	107227	TGTAATCCCAACACTTTGGAGAGCTTGAGCAAAAGAAATCACTTGAACCCAGAGGTTCAAG	107286
QY	540	ACTAGGCTTGCCCAACAATGGCGAAACCTGTCTCTAATAAAAAAATAGAAATTAAGCTGG	599
Db	107287	ACCAGGCTTGCTCAACATGGTGAATCTGTCTCTT-ACATAAAATACAAAATTTAGCCAGG	107345
QY	600	CATGTGGCGGGCGCCTGTATCTCACTGCTTCAGAGGCCGAGACACAAGATCACTTA	659
Db	107346	CATGTGGCGGCACACCTGTATCTCCAGTACTCAATGAGGTGAGGCACAAGATCACTTG	107405
QY	660	AACCCAGAGAGTGGAGGTTGCACTGAGCTGAATCTGTGCACCT-CACTCCAACTTGGGAG	718
Db	107406	AACCCGGGAAGCGAGGTTGCAGTGAAGCCGAGATCAAGCCACTGACTCCAGGCTTGTGTG	107465

QY	719	CAAGGTACACTTTTGCTCAAAAAAGAAAAAACAGTAAAAAAGAAACAGSTA	778
Db	107466	ACAGGCAAGACTCTGCTCAAAAAAATAAAAAATCAAGGAAAGAAACAATTGA	107525
QY	779	AGTTACTTTATTAACCAAGTATCCCAATACATCATTTCAAGTAAATTAATA	838
Db	107526	ATTATTTTACCAATTTTATTATTAACCAAGTACATCCAAAAGATTATCATTTCAAAATG	107589
QY	839	AAACATTATGATGATGACTTTACATTCCTTTCTGTTCATATTTAAGTCTTTGAAA	898
Db	107586	CAATTCATATTAATTAATTATGAGATATTTTAACTCTCTCTTTAATGCTATGCTTTGACA	107645
QY	899	GTGAGTATATGTTATGCTGACACACATCTCAATTTGAC	940
Db	107646	TCATGTGTGTGTTTACACTTACAGCCCAATTTTAATGGAC	107687

Search completed: February 27, 2005, 05:14:33
Job time : 318.935 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 13:59:12 ; Search time 1077.07 Seconds
(without alignments)
11008.461 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagagcccccag.....tcgaggggggaaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	252.6	12.6	52216	US-09-747-810-1	Sequence 1, Appl1
2	252.4	12.6	56098	US-10-087-192-1612	Sequence 1612, Ap
3	240.2	12.0	55827	US-09-813-133A-3	Sequence 3, Appl1
4	240.2	12.0	55827	US-10-212-877-3	Sequence 3, Appl1
5	239	11.9	1980090	US-10-719-993-6815	Sequence 6815, Ap
6	239	11.9	1980090	US-10-741-600-17676	Sequence 17676, A
7	235.6	11.8	27189	US-10-087-192-1828	Sequence 1828, Ap
8	235.6	11.8	84105	US-10-741-601-5637	Sequence 5637, Ap
9	229.2	11.5	653122	US-10-087-192-226	Sequence 226, Ap
10	227.4	11.4	23071	US-09-764-864-1673	Sequence 1673, Ap
11	226.2	11.3	133955	US-10-087-192-1984	Sequence 1984, Ap

12	225.4	11.3	325791	US-09-768-185A-1	Sequence 1, Appl1
13	225.2	11.3	35641	US-09-962-436-306	Sequence 306, Ap
14	225.2	11.2	35641	US-09-880-107-2225	Sequence 2225, Ap
15	224.8	11.2	629	US-10-027-632-252422	Sequence 252422, Ap
16	224.8	11.2	629	US-10-027-632-252423	Sequence 252423, Ap
17	224.8	11.2	629	US-10-027-632-252422	Sequence 252422, Ap
18	224.8	11.2	629	US-10-027-632-252423	Sequence 252423, Ap
19	224.8	11.2	629	US-10-681-818-155	Sequence 155, Ap
20	224.8	11.2	629	US-10-681-818-156	Sequence 156, Ap
21	221.6	11.1	2591	US-10-755-889-83	Sequence 83, Appl
22	221.2	11.1	84409	US-10-741-601-5696	Sequence 5696, Ap
23	221.2	11.1	84409	US-10-741-600-17771	Sequence 17771, A
24	220.8	11.0	66666	US-09-736-960-86	Sequence 86, Appl
25	220.8	11.0	96587	US-09-997-722-250	Sequence 250, Ap
26	220.6	11.0	203264	US-10-087-192-988	Sequence 988, Ap
27	220.2	11.0	126990	US-10-717-597-68	Sequence 98, Appl
28	220	11.0	32146	US-09-764-860-797	Sequence 797, Ap
29	220	11.0	32146	US-10-074-095-797	Sequence 797, Ap
30	220	11.0	32146	US-10-212-872-797	Sequence 797, Ap
31	220	11.0	32248	US-09-764-860-802	Sequence 802, Ap
32	220	11.0	32248	US-10-074-095-802	Sequence 802, Ap
33	220	11.0	32248	US-10-212-872-802	Sequence 802, Ap
34	220	11.0	96589	US-10-052-482-214	Sequence 214, Ap
35	219.2	11.0	100267	US-10-240-425-1470	Sequence 1470, Ap
36	218.6	10.9	4329	US-09-764-891-8960	Sequence 8960, Ap
37	218.6	10.9	43159	US-10-741-600-17853	Sequence 17853, A
38	218.6	10.9	166236	US-10-741-600-17759	Sequence 17759, A
39	218.6	10.9	366803	US-10-719-993-6805	Sequence 6805, Ap
40	218.4	10.9	2312	US-09-764-877-3606	Sequence 3606, Ap
41	218.4	10.9	2312	US-10-242-515-3606	Sequence 3606, Ap
42	218	10.9	2166	US-10-027-632-98060	Sequence 98060, A
43	218	10.9	2166	US-10-027-632-98060	Sequence 102722, A
44	218	10.9	2166	US-10-027-632-98060	Sequence 98060, A
45	218	10.9	2166	US-10-027-632-102722	Sequence 102722, A

ALIGNMENTS

RESULT 1
US-09-747-810-1
Sequence 1, Application US/09747810
Patent No. US20020012903A1
GENERAL INFORMATION: Goldman, Steven A.
APPLICANT: Goldman, Steven A.
TITLE OF INVENTION: A METHOD FOR ISOLATING AND PURIFYING MULTIPOTENTIAL NEURAL PROGE
FILE REFERENCE: 19603/3580
CURRENT APPLICATION NUMBER: US/09/747, 810
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/173,003
PRIOR FILING DATE: 1999-12-23
NUMBER OF SEQ ID NOS: 2
SEQ ID NO 1
LENGTH: 52216
TYPE: DNA
ORGANISM: Homo sapiens
US-09-747-810-1

Query Match 12.6% Score 252.6; DB 9; Length 52216;
Beet Local Similarity 69.1%; Pred. No. 2.4e-50;
Matches 403; Conservative 0; Mismatches 174; Indels 6; Gaps 4;

Qy 417 TTTCAGTTTATAGTACCACTTAACAGGTAACAGGCTGGCGAGTGGCTCAC 476
Db 41515 TTAATAATTTTCTAAGAACCACTTAA--AAGACATAAAGCGCGCGGTGCTCAC 41572
Qy 477 ACCTGTATCCAGCACTTTGGAGGCTGAGCGAGATCACTTTGTCTGAGATTT 536
Db 41573 TCCGTATATCCAGCACTTTGGAGGCTGAGCGAGATCACTTTGTCTGAGATTTG 41632
Qy 537 GAGACTAGCTGGCGAACATGGGAACTGTCTCTAAAAAATAAATAAATTAGCC 596

```
Db 41633 GAAACAGCCTGGCCAAACAGGGTGAACCATGCTCT-ACATAAAATACAAAATTAGCT 41691
QY 597 TGGCATGCTGGGGGGGCGCTGTATCTCAGCTGCTCAGAGGCGCGAGACAAGATCNC 656
Db 41692 GGGTGTGTGTGGTGGTCTGTATCGCAGCTCTCGGGAGGCTGAGCGAAGAAATCAT 41751
QY 657 TTAACCCAGAGGTGAGGTTGTCAGTGAAGTATCGTGCACCT-CACTCCAACTGG 715
Db 41752 TTGAAGGAAGAGGTGAGGTTGCATGAGCCAAAGATGGCCACTGCACTCCAGCTGG 41811
QY 716 GAGACAGAGTGAACCTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAGAAAAGG 775
Db 41812 GTAAAAAGATGACATCCATCTCAAAAAATAAATAATTAATTAATAAATAA 41871
QY 776 TGAAG--TTAACTTATATACCAATGATCCCAATACAAATCATTTCAAGTAA 833
Db 41872 AAAAGACATATAATGAACAGTGAATTTATTTATATATATATCAAAAAATTACGTT 41931
QY 834 ATATAAAACAATTATGATAGATTAATCTTACATTTCTTTCTTTTCAATATTAAGCTT 893
Db 41932 CAACATGATATCAATGATAAATTAATATATCACTGATATTTTCACTTTCTGCACTT 41991
QY 894 TGAAGTGAATATATATGTTATGCTGAACAGCAATCTCAATTTGAGTAACTTCA 953
Db 41992 TGATATCCATGATATATTTTGTGACCTTACAGCACTGGTATGTTGGCCAGCTCATCTCA 42051
QY 954 GGTGCTCAGTACGACATGAGCTAGAGTACTGATTTAGAT 996
Db 42052 AGTGCTCAGTACGACACGTTGGTGAAGTGTCACTTTATAGAT 42094
```

RESULT 2

```
US-10-087-192-1612/C
; Sequence 1612, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1612
; LENGTH: 56098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(56098)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1612
```

```
Query Match 12.6%; Score 252.4; DB 13; Length 56098;
Best Local Similarity 70.8%; Pred. No. 2.8e-50;
Matches 408; Conservative 0; Mismatches 156; Indels 12; Gaps 5;

QY 440 TTTAAACAGTAAAGGTGGGCGCAGTGGCTCACACTGTAATCCAGCACTTTGGG 499
Db 26680 TTTAAACAGTAAAGGTGGGCGCAGTGGCTCACACTGTAATCCAGCACTTTGGG 26621
QY 500 AGGCTGAGGAGGAGATCACTTTGGTGAAGTTGAGACTAGAGCTGGGCAACATGGG 559
Db 26620 AGGCTGAGGAGGAGATCACTTTGGTGAAGTTGAGACTAGAGCTGGGCAACATGGG 26561
QY 560 GAAACTCTGTCTTAAATAAATAATTAAGCTGAGTGTGGCGGGCGCTGTA 619
```

```
Db 26560 GAAACCCATCTCT-ACATAAAATAGAAAAGTTAGCCGATGTGGGACACGCTTGTGA 26502
QY 620 ATCTACGCTGCTCAGAGGCGCGAGACAAAGATCACTTAAACCAAGAGGTGGAGTTGG 679
Db 26501 GTCCAGACGCTCGGAGGCGAGACAAAGATCACTTAAACCGCAAGGTGGAGTTGG 26442
QY 680 CAGTGAAGTGAATGCTGCA-CTCACTTCAACTGGAGAGAGATGACACTTTGTCT 738
Db 26441 CAGTGAAGTGAATGCTGCACTTCACTTCACTTGGGAGAGAGAGATGACATCTC 26382
QY 739 CAAAAAGAAAAAACAAGTAAAAAAGAAACAGTGAAGTTAACTTTAATTAACCA 798
Db 26381 AAAAAAACAATAA-----CAAAACAAAAATGATATATTTAATAATTA 26328
QY 799 TGAATCCCAATATACATCTTCAAGTGAATTAATTAATAAACAATTATGAATGATGA 858
Db 26327 TTTATTTAAACCGTTATATCTAAATGTATGTAATGTAATGTAATGTAATTTTA 26268
QY 859 CTTTACATCTTTCTTTCTTTTCAAT-TTAAGCTTTGAAGTGAATATATATGTTATGC 917
Db 26267 ATGGAATATTTTACATCTTTTAAAGATATGTTGAAGTCACTGATATATACAC 26208
QY 918 TGAACAGACA--TTCATTTGAGTGAATCAATTTCAAGTGTCTCAGTACCAATGTG 974
Db 26207 TTAATGACAACTTCTCAATTTGAGTGAATTTCAAGTGTCTCAAGTGTCTCAAGTGTG 26148
QY 975 GCTAGAGTACTGATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1010
Db 26147 GCTAGAGTACTGATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 26112
```

RESULT 3

```
US-09-813-133A-3/C
; Sequence 3, Application US/09813133A
; Publication No. US20020137179A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001173
; CURRENT APPLICATION NUMBER: US/09/813,133A
; CURRENT FILING DATE: 2001-06-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Human
US-09-813-133A-3
```

```
Query Match 12.0%; Score 240.2; DB 9; Length 55827;
Best Local Similarity 68.6%; Pred. No. 2.7e-47;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCCTAGAGTGTGTTTGTAGTAATAATGCTGAGCTGTTATGTCATTTTCA 422
Db 46943 TGAATCTCAGACAGGCTTGTATTAAGAAATATATAGCAAGTCAACATATTTACTTAA 46884
QY 423 GTTTTCTAGACCACTTAAACAGTAAAC--AAGCTGGGCGCAGTGGCTACACCT 480
Db 46883 ATTTTCAAGTTACACTTTTAAGAAAAATTAAGAAAGGCGAGGCAAGGCTGACACCT 46824
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGCGAGCAAGTCACTTTGTCTGAGAGTTTGA 540
Db 46823 ATTAATCCAGCACTTTGGAGGCTGAGGCGAGGCGAGTCACTTGAAGTCTGAGAGTTTGA 46764
QY 541 CTAGCTGGGCAACATGGGAAACTCTGCTCTTAAAAAATAAATAATTAAGCTGGC 600
Db 46763 CCAAGGAGGCAACATGGGAAACCCGCTCTTACCAAAA--TACAAAAATTAAGCAAGC 46706
QY 601 ATGTGGGCGGCGCTGTATCTCAGTGTCTCAGAGGCGGAGACAAAGAAATCACTTAA 660
```



```

Db      46705 ATAGTGTGACGCGCTGTAATCCAGCTACTGTGGAGGCTGAGGACGAAACCTTGCTTGA 46646
QY      661 ACCGAGAGGTGGAGGTTGTCAGTGAAGTGTGCTCACT-CATTCGAACCTGGGAGA 719
Db      46645 ACCGAGAGGAGAGGTTGTCAGTGAAGTGTGCTCACT-CATTCGAACCTGGGAGA 46586
QY      720 CAGAGTGACA-CTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
Db      46585 CAGAACAGACCTCTGCCACACACACACACAAAAAGAAATATGTGAGATTAAATCAATA 46526
QY      779 AGTTAACTTTAATTAACCAATGATCCCAATATCAATCAATTTCAAGTGAATATATA 838
Db      46525 TTGTAATTTAATTAACCAATATGATCAAGTATCA-----TTTCAAGTGAACCAATATA 46470
QY      839 AAACAATTAATGAATGAGTACTTTCACTCTTTCTGTTTCAATTAATTAAGCTTTGAAA 898
Db      46469 AA--AATTATCTTGAATAATTTAATGTTACTT-----TTTCAATTAAGCTTTCAAT 46417
QY      899 GTGAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTAGTACATTTGAGGTGC 958
Db      46416 TCTGCTGTGTTAATTAACATCTCACTGATGTTCTCAATTCAGACAACTACATTTTAACGAC 46357
QY      959 TCAGTAGCCACATGTGCTAGCACTTACTGTATTGA 995
Db      46356 TGATAGCCACATGTGTTAGTGTGCTACTACTGATGGA 46320

```

```

RESULT 4
US-10-212-877-3/c
; Sequence 3, Application US/10212877
; Publication No. US20030017574A1
; GENERAL INFORMATION:
; APPLICANT: GAN, Weiniu et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL00117301V
; CURRENT APPLICATION NUMBER: US/10/212,877
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/813,133
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 55827
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-212-877-3

```

```

Query Match      12.0%; Score 240.2; DB 14; Length 55827;
Best Local Similarity 68.6%; Pred. No. 2.7e-47;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY      363 TCATACCTAGAGTAGTGTGTTAGTAAATATGCTGAGCTTATGTCATTTCGA 422
Db      46943 TGAATCTCTGACACAGAGGTGTTAATAGAAATATATGCAATCAATATTTACTTTAA 46884
QY      423 GTTTTATAGTACCATTAACAGGTAA--AAGCTGGGCGAGTGGCTCACACT 480
Db      46883 ATTTTCAAGTTACACTTTTAAGAAATAAAGAAAGGCGACGCGTGGCTCACACT 46824
QY      481 GTAATCCAGACATTTGGAGGCTGAGGCGAGATCACTTTGGTCAAGAGTTTGA 540
Db      46823 ATATATCCACACATTTGGAGGCTGAGGCGAGATCACTTTGATCAGAGTTTGA 46764
QY      541 CTAGCTGGGCAACATGGCGAACTGTCTCTTAAAAAAAATACAAATAATAGCTGGC 600
Db      46763 CCAAGGAGGCGCAATGTAATGTAATCCCGTCTCTTCAAAAA--TACAAAAATAGCCAGGC 46706
QY      601 ATGCTGGCGGCGCTGTAATCTTCACTGCTCAGAGGCGGAGACACAAGATCACTTAA 660
Db      46705 ATAGTGTGACGCGCTGTAATCCAGCTACTGTGGAGGCTGAGGACGAAACCTTGCTTGA 46646

```

```

QY      661 ACCGAGAGGTGGAGGTTGTCAGTGAAGTGTGCTCACT-CATTCGAACCTGGGAGA 719
Db      46645 ACCGAGAGGAGAGGTTGTCAGTGAAGTGAATGATGACATGCTGACCTCGAGGTGA 46586
QY      720 CAGAGTGACA-CTTTGTCTCAAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGA 778
Db      46585 CAGAACAGACCTCTGCCACACACACACACAAAAAGAAATATGTGAGATTAAATCAATA 46526
QY      779 AGTTAACTTTAATTAACCAATGATCCCAATATCAATCAATTTCAAGTGAATATATA 838
Db      46525 TTGTAATTTAATTAACCAATATGACAAAGTATCA-----TTTCAAGTGAACCAATATA 46470
QY      839 AAACAATTAATGAATGAGTACTTTCACTCTTTCTGTTTCAATTAATTAAGCTTTGAAA 898
Db      46469 AA--AATTATCTTGAATAATTTAATGTTACTT-----TTTCAATTAAGCTTTCAAT 46417
QY      899 GTGAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTAGTACATTTGAGGTGC 958
Db      46416 TCTGCTGTGTTAATTAACATCTCACTGATGTTCTCAATTCAGACAACTACATTTTAAGGAC 46357
QY      959 TCAGTAGCCACATGTGCTAGCACTTACTGTATTGA 995
Db      46356 TGATAGCCACATGTGTTAGTGTGCTACTACTGATGGA 46320

```

```

RESULT 5
US-10-719-993-6815/c
; Sequence 6815, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6815
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
US-10-719-993-6815

```

```

Query Match      11.9%; Score 239; DB 18; Length 1980090;
Best Local Similarity 72.6%; Pred. No. 2.9e-46;
Matches 416; Conservative 0; Mismatches 125; Indels 32; Gaps 7;

QY      437 ACAATTAACAGGTAAATAAAGGCTGGGCGCAGTGGCTCACACCTGTATCCAGACATT 496
Db      539204 AAATTAAGAAATATGAGAGGCTGGGCTGGTGTCTCATGCTGTATCCAGACATT 539145
QY      497 GGAAGCTGAGGAGGAGGAGATCACTTTGGTCAAGATTGAGACTAGCTGGCCAAAT 556
Db      539144 GGAAGGCGAAGGCGGAGGAGATCACTGAGGTCAAGATTGAGACAGGCTGGCCAAAT 539085
QY      557 GCGGAAATCTGTCTCTTAATAAATAAATAATTAAGCTGGCATGTGTGGCGGCGCT 616
Db      539084 GGTGAATCCCGTCTCT-ACTTAATAATGACAAATTAAGCGGCGGTGTGGCGGCT 539026
QY      617 GTAATCTAGCTGCTCAGAGGCGGAGACACAAGATCACTTAAACCCAGAGGTGAGG 676
Db      539025 GTAATCCAGCTACTCAGAGAG--CGAGCGAAGAAATGTCTCAAAACCCAGAGGTGAGG 538967
QY      677 TTGCACTGAGCTGAGATGCTGCTCACT-CATTCGAACCTGGGAGACAGAGTGAACCTTTG 735
Db      538966 TTGCACTGAGCTGAGATGCTGCTCACTCACTGAGCTGGGAAACAGAGAGATCTCAGT 538907

```

Oy	736	ICTCAAAAGAGAAAAAACAAGTAAAAAGAAACAGGTGAAGTTAACTTTAATAC	795
Db	538906	TGAAGAAAAAGAAACAGAGAAATTACAAATA-----TATTCAATTACT	538859
Oy	796	CAAGTATATCCCAATACATCTTCAAGGTAAATTAATATAAACAATTATGAATGAG	855
Db	538858	CAATGTATCCAAATTATGGTCATTTTAGCATGTAATCAATATATGAG-----TTTATATGAG	538804
Oy	856	ATACTTTCATCTTTCTGTGTTTCATTTAAGTCTTTGAAGTGAATATATATAT	915
Db	538803	ATAGTCTATGTTCTTTTTCCTAAG-----TCTGAATTTGGGTGTAT-TTAC	538756
Oy	916	GCTGACAGACATCTCAATTTTGGACTAGCTACATTTCAAGTGTCTAGTACCAATGTGG	975
Db	538755	ACTGACAGACATCTCAATTTTAGACTAGCCACATTTCCAGTGTGTGAGTGCCTCATGTAC	538696
Oy	976	CTAGCAGTTACTGTATTTGATGGCAGGATCTTA	1008
Db	538695	CTGGTGGCTTACCACTGGACAGCAACAGGTCTTA	538663

```

RESULT 6
US-10-741-600-17676/c
; Sequence 17676, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING NOS: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17676
; LENGTH: 1980090
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1980090)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-600-17676

```

Query Match	11.9%	Score 239;	DB 19;	Length 19809090;
Best Local Similarity	72.6%;	Pred. 2.9e-46;		
Matches 416;	Conservative 0;	Mismatches 125;	Indels 32;	Gaps 7;
Qy	437	ACATTAAACAGGTAAACAAAGGCTGGGCGCAGGCGTCAACCTGTAAATCCAGCACTTT	496	
Db	539204	AAATGAAAGAAATAGAGAGGCTGGGCGTGGGCTGATGCTGTAAATCCAGCACTTT	539145	
Qy	497	GGGAGGCTGAGGCGAGGAGATCACTTTGGTCAGAGATTGAGATCAGGCTGGCCAAAT	556	
Db	539114	GGGAGGCCAAGGCGGGCAGATCACTGAGGTGAGAGTTTGAGACCAAGCTGGCCAAAT	539085	
Qy	557	GGCGAACTCTGTCTCTTAAAAAAAATACAAAAATTAGCGTGGCATGTGCGCGGCGCT	616	
Db	539084	GGTGAAACCCCGTCTCT-ACTTAAAGTACAAAAATTAGCCGGCGGTGTGGCAGGTGCT	539026	
Qy	617	GTAATCTCAGCTGCTCAGGAGGCCGAGACACAAGATCACTTAAACCAGAGGTGGAG	676	
Db	539025	GTAATCTCCAGCTACTCAGGAGG-CGAGGCGAAGAAATTGCTCAAAACCCAGAGGTGGAAG	538967	
Qy	677	TTGCAGTGAGCTGAGATCGTGCACCT-CACTCCAACTTGGGAGACAGAGTGACACTTTTG	735	
Db	538966	TTGCAGTGAGCCAGATTGTGCCACTGCACTCCAGCGTGGGAAAAGAGGAGACCTCAGT	538907	
Qy	736	TCTCAAAAAGAAAAAAAACAAGTAAAAAAGAAAAGAGTGAAGTTAACTTTAATTAACC	795	
Db	538906	TGAAAAAAAAGAAACAGAGAAATTTACAAATA-----TATTTCAATTAACTC	538859	
Qy	796	CAATGTATCCCAATATCAATCTTCAAGTGAATTAATATTAACAATTTATGAATGAG	855	

Accession	Sequence	Position
Dd	538858 CAAATGATCCAAATATAGTGATATTTACAGTATCAATATAGA-----TTTAAATGAG	5388040
Oy	856 ATACTTACATCTCTTTCTGTGTTTCATATTAGCTTTGGAAGTGATATATGTAT	915
Dd	538803 ATAGCTATGTTCTTTTCTCTAAG-----TCTGAATTCGGTGTGTAT TTAC	5387566
Oy	916 GCTGCAGCAGCATCTCAATTGTAAGTACTACATTTCAAGTCTCAGTACCAATGTGG	975
Dd	538755 ACTGCAGCAGCATCTCAATTTTGAAGTACGCCACATTTCCAGTCTGAGTGGCTGCATGTAC	5386586
Oy	976 CTACAGTTACTGTTATTTGATGAGCGAGATCTA	1008
Dd	538695 CTGGTGGCTACCACTGGACAGCAGAGTCTA	538663

```

RESULT 7
US-10-087-192-1828/c
; Sequence 1828, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1828
; LENGTH: 27189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1828

```

	Query Match	11.8%;	Score 235.6;	DB 13;	Length 27189;
	Best Local Similarity	68.8%;	Pred. No.2,5e-46;		
	Matches 385;	Conservative 0;	Mismatches 164;	Indels 11;	Gaps 4;
QY	439	ATTAAACAGGTAAAAAAGCGTGGGCGCAGTGGCTCACACTTGAATCCACACTTTGG	498		
Db	11830	AGTAAATGTGCTAAAAATGGGCGGGCATGTATCTATCTTGTATCCACACTATGA	117711		
QY	499	GAGGCTGAGGCAGGCGATGATCCTTTGTGACGAGATTGAGACTGACCTGGCCACATGG	558		
Db	11770	GAGGCGAGGTTGGGCGAGATCCTTAGGCTCAGGAGTTTGAACCGAGCTGGCCACATGG	117111		
QY	559	CGAACTCTGTCTCTAAAAAAATACAAAAATTAGCCTGGCATGGTGGCGGCGCTGT	618		
Db	11710	TGAAACCCGCTCTCT-ACTAAGAATATAAAAATTAGCTGGAGTGGTGGTGGCTGT	116522		
QY	619	AATCTCAGCTGCTCAGAGGCGGAGCACAAATCACTTAAACCAGAGGTGAGGTT	678		
Db	11651	AATCTCAGCTACTCAAGAGCTGAGGCGAGAAATTTCTTGAACCCAGAGGCGAGAGTT	115922		
QY	679	GCAGTGAGCTGAGATGATGTGCCACTCACTCCAACTGGGAGACAGATGACACTTTTGTCT	738		
Db	11591	GCAGTGAGCGGAGACTGCACCAATTGCACTCCAGCATGATGGCAGAGGCGAAAAGTCTGT	115322		
QY	739	CAAAAAGAAAAAACAAGTAAAAAAGAAACAGGTGAAGTTAATTATTAACCCA	798		
Db	11531	CAAAAAAAGAAAAA-----AAAAAGCAACAGCTTTTGGCAGATGATCAACATTAC	114777		
QY	799	TGATCCCAAAATCATCTTTCAAGATGATTAATATATAAACAATTGTGATGAGATA	858		
Db	11476	TACACTCCCTTCTCCCTTCACTTTTCAAAAAAATTTTTCATGCATTTTAACAAACATTATTTA	114172		
QY	859	CTTACATCTCTTTCTGTGTTTCATATTAAGTCTTGAAGAGTGAAGTATATATGTTATGC-	917		

Db 11416 GATTTCATCTC-----TCTCTCATGTTAACTCTTTAAATCTGGGGTAAATTTTAACT 11366

QY 918 TGAACGACATCTCAATTTGAGCTAGCTACATTTCAAGTGTCTGATGACACATGTGGCT 977

Dd 11360 TAAAGACTGTCTTAATTTGGCATGACACATTTCAAGTGTCTGATGACACAGTGAT 11301

QY 978 AGCAGTACTGTATTGGATG 997

Dd 11300 AGTGTCACTGTATTAGTGG 11281

```

RESULT 8
US-10-741-601-5637/C
; Sequence 5637. Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMERISMS ASSOCIATED WITH
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5637
; LENGTH: 84105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-601-5637

```

Query Match	11.8%	Score	235.6	DB	18	Length	84105
Best Local Similarity	67.4%	Pred. No.	4.4e-46				
Matches	407	Conservative	0	Mismatches	189	Indels	8
						Gaps	5

Oy	362	CTCATACCCTTAGAGTAGGGGTTAACTAGAAAT--AACTGCTAGGCTGTATGCAATT	419
Db	68837	CTCAGAGCTTAGAGTAGCTGTGTCAGAGAAATACATTTAAACCACTATATTTATCT	68778
Oy	420	CCAGTTTTTTTAGTAGCCACTTAAACAGGTAAAAAAAGCTGGGCGCAGTGGCTCAACC	479
Db	68777	TAAATTTTCTAGTTGGCTCATTTAAAAAAGTAAAGAGACGGGCGGAGTGTCAATC	68718
Oy	480	TGTAAATCCCGACACTTTGGGAGGCTGAGGCGAGCAATCACTTTGGTCAGGAGTTTGAG	539
Db	68717	TGTAAATCCCGACACTTTGGGAGGCGGAGGCGGTGCAATCA--GAGGTCAAGGATCAAG	68660
Oy	540	ACTAGCCTGGCCACATGGCGAACTGTCTCTCAAAAAAAAATACAAAAATTAGGCTGG	599
Db	68659	ACCATCTGTGCTAACATGTGTAAACCCCGCTTTTCTAAAAATACAAAAATTAGCCAG	68600
Oy	600	CATGTGGCGGGCGCTGTAACTTCAGCTGCTCAGAGGCGGAGACACAGAAATCATTA	659
Db	68599	CGTGTGGCGAGGCACTGTAGTCCCACTGCTCGGGAGGCTGGGGCAGAGAAATCACTGG	68540
Oy	660	AACCCAGAGGTGAGGTGCAGAGGACTAGATCTGGCCACT--CACTCCACTGGGAG	718
Db	68539	AACCCGAGGCGAAGGTGTGAGTGAACCAAGATCGGCCACTGTCACTGACTGGGGTG	68480
Oy	719	ACAGAGTGACATT-TGTCTCAAAAAAGAAAAAAAACAACTAAAAAAAGAAACAGGTG	777
Db	68479	ACAGAGAGAGCTCATCTCAACCAACAAACAAACAAACAAATTAATAATTAAGAGA	68420
Oy	778	AAGTTAATCTTAATAACCCAACTGATCTCCAAATACATCACTTCAAAAGTAAATTAAT	837
Db	68419	TAGGTAGAAATGATTTTATGTTTTAT--CTPACCCAAATATATCAAAATATTTCAATTGA	68362
Oy	838	AAAACAATTATGAATGAGTACTTATCATCTTTCTGTTCATATTAAGCTTTGAA	897
Db	68361	ACATGTATCAATATAAAAAATTAAAGTACTTTATATCTTTTTCACACTTAAGATTGGA	68302
Oy	898	AGTGAATATATATGTTATGCTGACAGCACTTCGAATTTGGACTAGCTACATTCAAGGTG	957

Db 68301 AATTGGATATATTTTCACTTATAGCATATCTCAATTGAGATAGCCTCATAGGTGGTG 68242

Qy 958 CTCA 961

Db 68241 GGCA 68238

```

RESULT 9
US-10-087-192-226/C
; Sequence 226, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 226
; LENGTH: 653122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) -- (653122)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-226

```

Query Match	11.5%	Score 229.2;	DB 13,	Length 653122;
Best Local Similarity	69.7%	Pred. No. 4.3e-44;		
Matches 404; Conservative	0;	Mismatches 153;	Indels 23;	Gaps 6;

OY	440	TTAAACAGGTAAACAAAGGCTGGGGCAGTGGCTCACCTGTAAATCCAGACCTTTGGG	499
Db	360810	TTTTAAATAATCAATTTGGGCCAGGTGGCGGTGCTCATGTCTTAAATCCAGACCTTTGAA	360751
OY	500	AGCTGAGGCGAGCGAGATCACCTTGTGTGAGAGATTGAGACTAGCTGGCCAAACATGGC	559
Db	360750	AGGCGAGAGGCGGGGCAATCAC--GAGGTGAGAGATTGAGACCAAGCTGACCAACATGGT	360693
OY	560	GAAACTGTGTCTCTAAACAAATTAACAAATTAGCCTGGCANTGGTGGCGGGCGCTGTA	619
Db	360692	GAAACCCCGCTCT-ACTAAATATCAAAAATTACCGAGCGTGGTGGTGCACGCGCTGTA	360634
OY	620	ATTCAGTGTGTCAGGAGGCGGAGCACAAGAAATCACTAAACCAAGAGTGTGAGGTTG	679
Db	360633	ATCCGAGCACTTGGGAGGCTGGAGCAGAGAGATACCTTGATCCCTGGGAGGAGAGGTTG	360574
OY	680	CAGTGAAGTGAAGATGTGCCACT--CACTCCAACTGGGAGCAGAGTGAACCTTTGTCT	738
Db	360573	CAATGAATGAGATTATATCCACTGACCTCCAGCCTGGGCGACAGAGCAAGACTTCATCTG	360514
OY	739	CAAAAGAAAAAACAAGTAAAAAGAAACGTTGAAGTTAACTTTAATAAACCA	798
Db	360513	AAAAAAAAAAAAAAAAAAAAAAAAAATTGG-----ATTTTACTTAA	360466
OY	799	TGTATCCCAATACATCATTTCCAAAGTAAATTAATPAACAAATTATGATAGATGA	858
Db	360465	TATATCCAAATAGTTTAAATTCAACATGACCATATATATAAAGATGATGATGATGA	360406
OY	859	CTTATACATCTTTTCTTGTTCATATTAAGCTCTTGAAGGATATATATATATGTCT	918
Db	360405	ACCTTCTTCTCTTTTTC---TACTAAGCTTTGATGTCACACATATATTTATACA	360349
OY	919	GACAGCAATCTCAATTTGAGTAACT---ACATTTCAAGTGTCTCAAGTACCAACATGTG	974

Db 360348 TATGCCATCTTGAATGGGGCTAGCTAGCCAGCTTCCAGGGCTCAAGGCCACATCTG 360289
QY 975 GCTAGCAGTACTGTATGTGATGCAAGGATCTAGAGGGA 1014
Db 360288 GCTGTAGCTACCATAGCAAGAGTGAAGGACCA 360249

RESULT 10
US-09-764-864-1673/C
; Sequence 1673, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION DATA REMOVED - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1673
; LENGTH: 23071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1673

Query Match 11.4%; Score 227.4; DB 9; Length 23071;
Best Local Similarity 70.1%; Pred. No. 2.4e-44;
Matches 445; Conservative 0; Mismatches 156; Indels 34; Gaps 9;

QY 381 GTGTTAGTGAATAATGCTGAGCTGCTATGTCATTTCCAGTTTCTTTAGTACCAAT 440
Db 10748 GTGCTAAGGAATATATATGATGAGCCATATATATTTAAATTTCTGCTCCACAT 10689
QY 441 TAAAAAG-GTAAAAAGCTGGCGCAGTGGCTCACTGTAATCCAGCATTTGG 499
Db 10688 AACAAAAAGTAAAAAGAGGCTGGCGCGGTGCTCATGCTGTATATCCAGCATTTGG 10629
QY 500 AGGCTAGGAGGAGCATACCTTTGGTCAAGTTTGAAGTCTGGCCAAACATGCG 559
Db 10628 AGGCTAGAGAGGAGGATCTGAGGCTGGGAGTTGAGACTTAACTAAACATGGA 10569
QY 560 GAAACTCTGTCT--CTAAAAAAAATACAAAAATTTAGCTGAGTGGCGGCGCT 616
Db 10568 GAAATCCATCTCTACTTAAAAAAAATTTCAAAATTTAGTGGCGGTGGGCAATGCT 10509
QY 617 GTATCTCACTGCTCAAGAGGCGGAGCAAGAACTACTTAAACCCAGAGGTGAG 676
Db 10508 GTATCTCACTGCTCAAGAGGCGGAGGCTGAGGAAATGCTTGAACCGGAGCGAGG 10449
QY 677 TTGAGTGAAGTGAATCGTGCACCT-CACTCCAGCTGGGAGACAGATGACATTTTG 735
Db 10448 TTGCGGTGAGCCAGATTTGCGCATTTGCACTCCAGCTGGGCAACAGCGAACTCCA 10389
QY 736 TCTCAAAAAAGAAAAAAAACAAGTAA-----AAAGAAACAGGTGAAGTTAACT 788
Db 10388 TCTCAAAAAAGAAAAAAAACAAGTAAAGTGAAGGAGGAGGAAAGTTAATTA 10329
QY 789 AA-----TAACCAATGATCCAAATACATATTTCAAGTGTATTAATTAATAAC 843
Db 10328 AATAGTGTAAACCAAGTATCCATGATATATTTCAACATGATATCAACATTA--A 10271
QY 844 AATATGATGAGATCTTATCATTTCTTTTCTTTTCAATTAAGTCTTTGAAGTGA 903
Db 10270 ACTATTTAGTGAATTTTACATTTGTT-----TATCATTTCTCAAAATCCAT 10223
QY 904 TATATATGTTATGCTG-ACAGCAGATCTCAATTTGAGTATACATTTCAAGGTCTCAG 962
Db 10222 TGTGTATTTTCTATCTTCAAGCAGCAACTCAATTTGGG--CAGCAGATTTCAATTTCTCA 10165
QY 963 TAGCCAGATGTGGCTGACGATTACTGTATTTGATG 997
Db 10164 TAGCCAGAAATGCTATTTGGCTATTTTATTTGTTG 10130

RESULT 11
US-10-087-192-1984
; Sequence 1984, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1984
; LENGTH: 133955
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1984

Query Match 11.3%; Score 226.2; DB 13; Length 133955;
Best Local Similarity 68.4%; Pred. No. 1.1e-43;
Matches 357; Conservative 0; Mismatches 158; Indels 7; Gaps 3;

QY 425 TTTTATAGTACCATTAATAACAGTAA-----AAAGGCTGGGCGAGTGGCTCACACC 479
Db 115167 TTTTATAGTACCATTTTAAACATTAAGAAAGAAACAGGCGGTGGTCTATGCC 115226
QY 480 TGTATCCAGCATTTTGGAGGCTGAGGAGGAGCATCACTTTGGTGAAGATTGAG 539
Db 115227 TGTATCCAGCATTTTGGAGGCTGAGGAGGAGCATCACTTTGAGCCAGAGTTCAAG 115286
QY 540 ACTAGCTGGCCAAATGCGGAACTGTCTCTAATAAAAAATCAAAATTTAGCTTGG 599
Db 115287 ACCAGCTGGGTAAACATGAGTGAATCTGTCTCT-ACATAAAATCAAAATTTAGCCAG 115345
QY 600 CATGTGGGCGGCGCTGTATCTGATCTGCTCAGAGGCGGAGACACAAGATCACTTA 659
Db 115346 CATGTGGGCGGACACTGTATCTCCAGTACTCAAGTATGATGAGGAGCAAGATCACTTG 115405
QY 660 AACCCAGAGGTGAGGTTGCAAGTGAAGTGAAGTGTGCACT-CACTCAACTGGAG 718
Db 115406 AACCCGAGAGGAGGTTGCAAGTGAAGTGAAGTGAAGTGTGCACTGTGTG 115465
QY 719 ACAGAGTGAATTTTGTCTCAAAAAAGAAAAAAAACAAGTAAAAAAAGAAACAGTGA 778
Db 115466 ACAGAGTGAATTTTGTCTCAAAAAAGAAAAAAAACAAGTAAAAAAAGAAACAGTGA 115525
QY 779 AGTTACTTTAATAACCAATGATCCCAATATACATCTTTCAAGTGTATTAATTA 838
Db 115526 ATTATTTTACCAATTTTATTTAACCAGATATCCAAAGATATATCATTTCAAAAG 115585
QY 839 AAACATTAATGAATGATATCTTACATTTCTTTTCTTTTCAATTAAGTCTTTGAAA 898
Db 115586 CAATTCATTAATTAATTAATTAATTAATTTTACATTTCTTTTCAATTAAGTCTTTGAAA 115645
QY 899 GTGAGTATATGTTATGCTGACAGACATCTCAATTTGGAC 940
Db 115646 TCAATGTGTGTTTACATTTACAGCCCAATTTTAATTTGGAC 115687

RESULT 12
US-09-768-185A-1
; Sequence 1, Application US/09768185A
; Publication No. US20040185439A1
; GENERAL INFORMATION:
; APPLICANT: Casseel, Michael et al

TITLE OF INVENTION: Betrogen receptor beta variants and
TITLE OF INVENTION: methods of detection thereof
FILE REFERENCE: CLO00280
CURRENT APPLICATION NUMBER: US/09/768,185A
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 09768185
PRIOR FILING DATE: 2001-01-24
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 325791
TYPE: DNA
ORGANISM: HUMAN
US-09-768-185A-1

Query Match 11.3%; Score 225.4; DB 11; Length 325791;
Best Local Similarity 82.6%; Pred. No. 2.6e-43;
Matches 270; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

Qy 441 TAAACAGGTAAAGGCTGGGCGAGTGGCTCAACCTGTATCCAGACCTTTGGGA 500
Db 282763 TGAAGAGAAACAGAGAGGCGGCGGCGAGTGGCTCAACCTGTATCCAGACCTTTGGGA 282822
Qy 501 GGGTGGGCGAGGAGATCACTTTGGTTCAGAGATTGAGACTAGCTGGCCAGATGCG 560
Db 282823 GGGCAAGGCGAGTGGATCATCTGAGTCAAGATTCAAGACTAGCTGGCCAGATGCG 282882
Qy 561 AAATCTGTCTCTTAAATAAATAATTAAGCTGGCATGTGGCGGCGCTGTAA 620
Db 282883 AAATCTGTCTCTCTTAAATAAATAATTAAGCTGGCATGTGGCGGCGCTGTAA 282942
Qy 621 TCTCAGCTGCTCAGAGGCGGAGACAGAGATCACTTAACCCAGAGGTTGAGGTTGC 680
Db 282943 TCCAGCTGCTCGGAGGCTTGAAGAGAGATGCTTGAACCCAGAGGTTGAGGTTGC 283002
Qy 681 AGTAGCTGAGATGCTGCACT-CACTTCAACTGGGAGAGAGATGACATTTGTCTC 739
Db 283003 AGTAGCTGAGATGCACTGCACTGCACTGCGGAGAGATGAGATCTGTCTCA 283062
Qy 740 AAAAAAAAAAAAAAAAACACTAAAAA 766
Db 283063 AAAAAAAAAAAAAAAAAAGAAAAA 283089

RESULT 13

US-09-962-436-306
Sequence 306, Application US/09962436
Patent No. US20020081301A1
GENERAL INFORMATION:
APPLICANT: Soppet, Daniel
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-75
CURRENT APPLICATION NUMBER: US/09/962,436
CURRENT FILING DATE: 2001-09-25
PRIOR APPLICATION NUMBER: US/60/235,082
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/234,924
PRIOR FILING DATE: 2000-09-25
NUMBER OF SEQ ID NOS: 568
SOFTWARE: PatentIn version 3.0
SEQ ID NO 306
LENGTH: 35641
TYPE: DNA
ORGANISM: Homo sapiens
US-09-962-436-306

Query Match 11.3%; Score 225.2; DB 9; Length 35641;
Best Local Similarity 84.4%; Pred. No. 1e-43;
Matches 265; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Qy 455 AAGGCTGGGCGAGTGGCTCAACCTGTATCCAGACCTTTGGAGGCTGAGGCGACA 514
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTCTGCTCAAAAAAAAAAAAAA 14988

Db 14689 AAGGCTGGGCTCAGTGGCTCAGCTGTATCCAGACCTTTGGAGGCTGAGGCGCC 14748
Qy 515 GATCACTTTGGTCAAGAGTTTGAAGTACTACCTGGCCAAATAGCCAAATCTGTCTCA 574
Db 14749 GATCACTTGAAGTCAAGAGTTTGAAGTACTACCTGGCCAAATAGTGAACCCGCTCTCA 14808
Qy 575 AAAAAAAAAATCAAAATTTAGCTGGCATGTGGCGGGCGCTGTATCTCAGTGTCTCAG 634
Db 14809 TTAAATAATACCAAAATTTAGCTGGCATGTGGCGGGCGCTGTATCTCAGTGTCTCAG 14868
Qy 635 GAGGCCGAGACACAGAAATCACTTAAACCCAGAGGTGAGGTTGCAGTGAAGTCAATC 694
Db 14869 GAGGCTGAGGAGAGAAATCACTTGAACCCAGAGGCGGAGGTTGCAGTGAAGTCAATC 14928
Qy 695 GTGGCACT-CACTTCAACTGGGAGAGAGTGAACACTTTGTCTCAAAAAAAAAAAAA 753
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTCTGCTCTCAAAAAAAAAAAAA 14988
Qy 754 AAACAGTAAAAA 767
Db 14989 AAAAAAAAACAAAAAA 15002

RESULT 14

US-09-880-107-2225
Sequence 2225, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-MO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2225
LENGTH: 35641
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L76568
US-09-880-107-2225

Query Match 11.3%; Score 225.2; DB 9; Length 35641;
Best Local Similarity 84.4%; Pred. No. 1e-43;
Matches 265; Conservative 0; Mismatches 48; Indels 1; Gaps 1;

Qy 455 AAGGCTGGGCGAGTGGCTCAACCTGTATCCAGACCTTTGGAGGCTGAGGCGACA 514
Db 14689 AAGGCTGGGCTCAGTGGCTCAGCTGTATCCAGACCTTTGGAGGCTGAGGCGACA 14748
Qy 515 GATCACTTTGGTCAAGAGTTTGAAGTACTAGCTGGCCAAATAGCCAAATCTGTCTCA 574
Db 14749 GATCACTTGAAGTCAAGAGTTTGAAGTACTAGCTGGCCAAATAGTGAACCCGCTCTCA 14808
Qy 575 AAAAAAAAAATCAAAATTTAGCTGGCATGTGGCGGGCGCTGTATCTCAGTGTCTCAG 634
Db 14809 TTAAATAATACCAAAATTTAGCTGGCATGTGGCGGGCGCTGTATCTCAGTGTCTCAG 14868
Qy 635 GAGGCCGAGACACAGAAATCACTTAAACCCAGAGGTGAGGTTGCAGTGAAGTCAATC 694
Db 14869 GAGGCTGAGGAGAGAAATCACTTGAACCCAGAGGCGGAGGTTGCAGTGAAGTCAATC 14928
Qy 695 GTGGCACT-CACTTCAACTGGGAGAGAGTGAACACTTTGTCTCAAAAAAAAAAAAA 753
Db 14929 AACTACTGCACTCCAGCTGGGCAACAGAGTGAAGTCTGCTCAAAAAAAAAAAAAA 14988

OY 754 AACACAGTAAAA 767
 Db 14989 AAAAAACAAAA 15002

Db 597 AAAAAAGAAACAGGTGAATTAATTTAAT 628

Search completed: February 27, 2005, 04:52:27
 Job time : 1097.07 secs

RESULT 15

US-10-027-632-252422
 ; Sequence 252422, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 252422
 ; LENGTH: 629
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-252422

Query Match : 11.2%; Score 224.8; DB 13; Length 629;

Best Local Similarity 72.8%; Pred. No. 1.8e-44;
 Matches 329; Conservative 1; Mismatches 118; Indels 4; Gaps 3;

OY 340 CAGGACCCCTGTTCTCAGAGCTCACTAGAGTGTGTTTATAGTAAATATG 399
 Db 181 CAGTACGTTTACATGAAATATATTAATTAAGGACGACATCCACAGAAATGAA 240
 OY 400 CTGAGCTGCTTATGTCAATTTCCAGTTTTTTATAGCCACATTAACACGATAAAAAAGGC 459
 Db 241 TACAGGCCACACATGCAATTTTAATTTTCTGTAGCCATACTTAATAAGTAAAGAGGC 300
 OY 460 TGGGCGGAGTGGCTCACACTGTATATCCGACACTTTGGAGGCTGAGGCGACATCA 519
 Db 301 TGGGTGAGTGGCTCAAGCTGTAAATCCAGAACTTTGGAGGCGCAAGGCGGATCA 360
 OY 520 CTTTGGTCAGAGTTTGAAGTCTGAGCTGCGCAACATGGGAAACTGTCTCTAAAAAA 579
 Db 361 --TGAGGTCAAGAGATGAACTCTGCGCAATATGTGAATCCCGTCTCT-ACTAA 417
 OY 580 AAAATACAAAAATTAGCTGAGTGTGCGGCGGCGCTGTATCTCAAGCTGCTCAGAGAGC 639
 Db 418 AAATACAAAAATTAGCTGAGTGTGCGGCGGCGCTGTATCTCAAGCTGCTCAGAGAGC 477
 OY 640 CGAGACACAGATCACTTAAACCCAGAGAGTGAAGTTGCACTGAGCTGAGATCGTGCC 699
 Db 478 TGAGGCGAAGAAATCGCTTAAACCCAGAGAGTGAAGTTGCACTGAGCTGAGATCGTGCC 537
 OY 700 ACTCATTCGAACCTGGAGACAGAGTGAACCTTTTGTCTCAAAAAAGAAAAAAACAA 759
 Db 538 ACTGCACTCCAGCTGATGTGATATGCGAGAC-TCATCTCAAAAAAAAGAAAAAGCA 596
 OY 760 GTAAAAAGAAACAGGTGAAGTTAATTTAAT 791

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 26, 2005, 20:26:59 ; Search time 6271.68 Seconds
(without alignments)
12144.524 Million cell updates/sec

Title: US-09-936-271b-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagaagcccccag.....tcgcaggggggaaacttta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g881.*
9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	11.3	908	6	CD245375
2	223	11.1	694	8	B2603010 WHAAX35TF
3	221.6	11.1	1750	3	AF119908 Homo sapi
4	220	11.0	769	6	CD370173 UI-H-FP1-
5	219.6	11.0	500	5	BX491983 DKFZP7810
6	218.4	10.9	837	2	BE790769
7	216	10.8	649	8	B71494
8	216	10.8	1019	6	CD518470
9	215.6	10.8	2449	3	CR749224 Homo sapi
10	215.4	10.8	573	5	BK477553 DKFZP686H
11	215.4	10.8	823	6	CD101954 AGENCOURT
12	215	10.7	397	1	AV759632
13	215	10.7	387	5	BK644875
14	215	10.7	633	7	CK905714
15	214	10.7	491	2	AM516097
16	214	10.7	4593	3	HM803561
17	213.4	10.7	517	8	AQ264959
18	213	10.6	569	8	AQ490878
19	212.8	10.6	700	5	BK508919
20	212.6	10.6	790	6	CD643827
21	212.4	10.6	619	5	BK485408
22	211.8	10.6	576	6	CD652305
23	211.8	10.6	731	7	CN307840
24	211.8	10.6	15970	8	AQ839852

25	211.4	10.6	537	8	AQ390925
26	211.4	10.6	624	9	AG019812 Homo sapi
27	211.4	10.6	654	7	CK003802
28	211.4	10.6	687	6	CD237038
29	211.4	10.6	692	5	B0662309
30	211.2	10.6	338	2	AM502796
31	211.2	10.6	596	8	AQ320557
32	211.2	10.6	801	8	B2600703
33	211	10.5	552	8	B68316
34	210.8	10.5	657	7	CF123536
35	210.6	10.5	1016	5	B0674146
36	210.4	10.5	388	2	AM069227
37	210.4	10.5	735	7	CF146929
38	210.2	10.5	671	9	AG079103
39	210.2	10.5	849	6	CB991080
40	209.8	10.5	397	1	A1421950
41	209.8	10.5	463	2	AM026305
42	209.6	10.5	461	8	B65075
43	209.6	10.5	501	5	BK485916
44	209.6	10.5	604	5	BM990713
45	209.2	10.5	534	6	CB050438

ALIGNMENTS

RESULT 1
CD245375/c
LOCUS
DEFINITION
AGENCOURT 14128276 NIH MGC 181 Homo sapiens cDNA clone
IMAGE:30374350 5', mRNA sequence.
CD245375
VERSION
CD245375.1 GI:31005839
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 908)
NIH-MGC <http://mgs.nci.nih.gov/>.
TITLES
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapds-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
<http://image.llnl.gov>
Plate: NDAM433 row: j column: 23
High quality sequence start: 10
High quality sequence stop: 557.
Location/Qualifiers
source
1. 908
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30374350"
/tissue_type="white Matter"
/dev_stage="Unknown"
/lab_host="DH10B-Ton A (T1 and T5 phage resistancees)"
/note="Vector: PCMV-SpOrt6.1; Site 1: NciI; Site 2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (copy site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 11.3%; Score 227; DB 6; Length 908;
Best Local Similarity 69.2%; Pred. No. 9.9e-26;


```
/clone="FLC0974"
/tissue_type="liver"
/dev_stage="fetus"
379..677
repeat_region
/rpc_family="Alu"
/rpc_type=dispersed
/complement(1163..1396)
/notes="predicted protein of HQ2955"
CDS
/codon_start=1
/product="PRO2955"
/protein_id="AA069662.1"
/db_xref="GI:7770253"
/translation="MPNCLILGRCSSFLILLIPYCSVLVLSGRSSCWPEKSKQPEV
KPKCRIMLVILILIKRKPAEMLVLVNINPYL"
ORIGIN
Query Match      11.1%; Score 221.6; DB 3; Length 1750;
Best Local Similarity 80.7%; Pred. No. 5.7e-25;
Matches 271; Conservative 0; Mismatches 64; Indels 1; Gaps 1;
QY 437 ACATTAAACAGTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 496
    |||||
DB 359 ACACGAAGGTGGGGAATATGGCCGGCGAGTGGCTCACGCTGTATCCAGCACTTT 418
    |||||
QY 497 GGGAGGCTGAGGCGAGATCACTTTGTCAAGAGTTGAGCTAGCCTGGCCACAT 556
    |||||
DB 419 GGGAGGCGGAGGCGAGATCACTTGAAGAGTTGAGCAAGCTGACCAAT 478
    |||||
QY 557 GGGCAAACTGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 616
    |||||
DB 479 GGTAAACCTGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 538
    |||||
QY 617 GTAATCTAGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 676
    |||||
DB 539 GTAATCCAGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 598
    |||||
QY 677 TTGAGAGAGTGTGATCGGCCACT-CATCTCAACCTGGGAGCAGAGTGAACCTTTG 735
    |||||
DB 599 TTGAGAGAGTGTGATCGGCCACTCGCTGACTTCACTGGGGGAGCAGAGAGTCTGT 658
    |||||
QY 736 TCTCAAAAAGAAAAAACAAGTAAAAAGAAA 771
    |||||
DB 659 CTCAAAAAGAAAAAAGATGAGAAATTAATCA 694
    |||||
RESULT 4
CD370173      769 bp      mRNA      linear      EST 05-AUG-2004
LOCUS         UT-H-FT1-bkb-e-05-0-UI.s1 NCI_CGAP_FTI Homo sapiens cDNA clone
DEFINITION    UT-H-FT1-bkb-e-05-0-UI 3', mRNA sequence.
ACCESSION     CD370173
VERSION       CD370173.1 GI:31154263
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 769)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-rt@mail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/cgap.html
The following repetitive elements were found in this cDNA
sequence: 346-636, >Alu
Seq primer: M13 FORWARD
```

```
POLYA=Yes.
FEATURES
SOURCE
Location/Qualifiers
1..769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT1-bkb-e-05-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: Score 1; Site 2: Not 1;
NCI_CGAP_FTI is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The RNA samples
were a mixture of these conditions (lines refer to
incubations following isolation by bronchoalveolar lavage)
(some normal donor macrophages were cultured in some of
the conditions, other donor macrophages in different
conditions). The mRNA samples were pooled for library
construction. Control 0 hours; control 3 hours; control 24
hours; LPS 100 ng/ml, 3 hours; LPS 100 ng/ml, 24 hours;
PMA 10 ng/ml, 3 hours; PMA 10 ng/ml, 24 hours; Klebsiella
moi 10, 3 hours; Klebsiella moi 10, 24 hours; Staph aureus
moi 10, 3 hours; Staph aureus moi 10, 24 hours; Adenoviral
vector (Ad5 CMV egfp), moi 500, 3 hours; Adenoviral vector
(Ad5 CMV egfp), moi 500, 24 hours; wt adenovirus moi 500,
3 hours; wt adenovirus moi 500, 24 hours; Ad vector + LPS
3 hours; Ad vector + LPS 24 hours; wt adenovirus + LPS 3
hours; wt adenovirus + LPS 24 hours. The library was
normalized according to Bonaldo, Lemon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
GGCCATGCGG. The tissue was provided by Dr. Gary W.
Hunninghake of the University of Iowa.
TAG TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UT-H-FT1
TAG_SEQ=GGCCATGCGG"
ORIGIN
Query Match      11.0%; Score 220; DB 6; Length 769;
Best Local Similarity 80.4%; Pred. No. 1.3e-24;
Matches 270; Conservative 0; Mismatches 65; Indels 1; Gaps 1;
QY 437 ACATTAAACAGTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 496
    |||||
DB 326 ACACGAAGGTGGGGAATATGGCCGGCGAGTGGCTCACGCTGTATCCAGCACTTT 385
    |||||
QY 497 GGGAGGCTGAGGCGAGATCACTTTGTCAAGAGTTGAGCTAGCCTGGCCACAT 556
    |||||
DB 386 GGGAGGCGGAGGCGAGTGTATCACTGAGGCCGAGATTGAACAGACCTGACAAAT 445
    |||||
QY 557 GGGCAAACTGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 616
    |||||
DB 446 GGTAAACCTGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 505
    |||||
QY 617 GTAATCTAGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 676
    |||||
DB 506 GTAATCCAGTCTCTTAAAGGCTGGGCGAGTGGCTCACTGTATCCAGCACTTT 565
    |||||
QY 677 TTGAGAGAGTGTGATCGGCCACT-CATCTCAACCTGGGAGCAGAGTGAACCTTTG 735
    |||||
DB 566 TTGAGAGAGTGTGATCGGCCACTCGCTGACTTCACTGGGGGAGCAGAGAGTCTGT 625
    |||||
QY 736 TCTCAAAAAGAAAAAACAAGTAAAAAGAAA 771
    |||||
```

Db 626 CTCAAAAAAGATGAGAAATATCA 661

RESULT 5
BX491983 500 bp mRNA linear EST 04-SEP-2003
LOCUS DKEZp78100513.1 781 (synonym: h1cc4) Homo sapiens cDNA clone
DEFINITION DKEZp78100513.5', mRNA sequence.
ACCESSION BX491983
VERSION BX491983.1 GI:32002977
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., Oesanger,A.,
Fodor,G., Han,W. and Wiemann,S.
TITLE EST (Koehler,K., Beyer,A., Mewes,H.W., Weil,B., Amid,C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMPZ (Biomedical Research Center at the Heinrich-
Heine-University, Dueseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No 81 sequence available.
This clone (DKEZp78100513) is available at the RZPD in Berlin.
Please contact the RZPD: Resourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: source@rzpd.de.

FEATURES
source
1..500
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKEZp78100513"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="781 (synonym: h1cc4)"
/note="Vector: pSport1_Sfi; Site_1: Sfi1A; Site_2: Sfi1B;
cDNA-collection"

ORIGIN

Query Match 11.0%; Score 219.6; DB 5; Length 500;
Best Local Similarity 82.7%; Pred. No. 1.7e-24;
Matches 263; Conservative 0; Mismatches 54; Indels 1; Gaps 1;

Dy 443 AAACAGGTAAAAAGGCTGGCGCAGTGGCTCAACCTGTATCCAGCACTTTGGAGG 502
Db 13 ATAAAGAAATGAGCGTGGGCAATGAGGCTCAACCTGTATCCAGCACTTTGGAGG 72
Dy 503 CTGAGGCGAGCATACCTTTGGTCTGAGAGTTTGAAGTACCTGCGCAACATGGCGAA 562
Db 73 CCGAGGGGGCGAGTCACTGAGGTCAAGGCTTCGAACTAGCCTGCGCAACATGGCGAA 132
Dy 563 ACTGTGCTCTAAATAAATTAAGCTGCGCATAGTGGCGGGCGCTGTATC 622
Db 133 ACCCATCTTTTCTAAATAATGCAAAATTAAGCGGGGTGTGTGTGAAGCGCTGTATC 192
Dy 623 TCAGTCTCTGAGAGCGCGAGACACAAGATCACTTAAACCCAGAGGTGAAGTTGCGAG 682
Db 193 CAGCTACTTGGGAGGGCTGAAGCAGAGAAATCGTTAAGTGAAGTGAAGTTGCGAG 252
Dy 683 TGAGCTGAATCGTCCACT-CACTCCAACCTGGAGACAGAGTGACACTTTGTCTCAA 741
Db 253 TGAGCTGAATGTTGGCAGCTGCACTTAAGCCTGGGTGAAGAGTGAAGTCTTGTCTCAA 312
Dy 742 AAAGAAAAAATCAAA 759
Db 313 AAAAAAAAAAAAAAAAAA 330

RESULT 6
BE790769/c 837 bp mRNA linear EST 20-SEP-2000
LOCUS 601581950F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3936325 5',
DEFINITION mRNA sequence.
ACCESSION BE790769
VERSION BE790769.1 GI:10211967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 837)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW780 row: c column: 14
High quality sequence stop: 723.

FEATURES
source
1..837
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936325"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_7"
/note="Organ: lung; Vector: pOT87; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAACAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 10.9%; Score 218.4; DB 2; Length 837;
Best Local Similarity 70.0%; Pred. No. 2.2e-24;
Matches 308; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

Dy 332 AAGACAGTGAAGACCCCTGTCTCAACAGAGCTCATCCCTAGAGTGTGTAGTAG 391
Db 472 AAAAAAGTAAGACCTTTATGGGCTGAATCATGTCTGCTGAATTCCTATGTTGAAG 413
Dy 392 AATAATGCTGAGCTGCTTATGCTATTCAGTTTATAGTGAAGCCATTAAGACGTA 451
Db 412 CCTTAACCCCGAGTACTGAAATGATGATTTGATATAGAGGCCCTTTAAAGGGTA 353
Dy 452 AAAAGGCTGGGGCGAGTGGCTCAACCTGTATCCAGCACTTTGGAGGCTGAGCGAG 511
Db 352 ATTAAAGCTGGGTGAGTGGCTCACACTATATCCCATCTTTGGAGGCGGAGCGAG 293
Dy 512 GCAGATCACTTTGTGTCAGAGTTTGAAGCTGAGCGGCGCAATGCGAACTGTCT 571
Db 292 GTGTATCAGCTTGAAGTCAGAGTTTGAAGCAGCTGAGCAACATTAAGAAACCTGTCT 233
Dy 572 CTAAAAAATAAATAAATTAAGCTGAGTGTGGCGGCGCTGTATTCAGCTGCT 631
Db 232 CTACTTAAAAATACAAAATTAAGCTTGTGTGTGTATACCGCTTATTCAGCACT 173
Dy 632 CAGAGGCGGAGACACAAGATCACTTAACCCAGAGGTGAAGTTGAGTGAAGTGAAG 691

[illegible]

OY	493	CTTTGGAGGCTGAGGCAGGCAGATCACTTGTGGCAGAGTTTAGACTACCTGGCCA	552
Dd	281	CTTTGGAGGCTTAGGCCAAGTGATCACTTAGGTCAGAAGTTCCAAACAGCTGGCCA	340
OY	553	ACATGGCGAAACTCTGTCTCTTAATAAAAAATTAACAAAATTAGCTGCACATGGTGGCGGC	612
Dd	341	ACATGGTAAACCCTGCTCT-ACATAAAATTAACAAAATTAGCTGCACATGGTGGCGGAC	399
OY	613	GCTGTAAATCTCAGCTGCTCAGAGAGCCGAGACAACAAGATCACTTAAACCCAGAGGTG	672
Dd	400	ACCTGTAATCTCAGTTATTTCAGGAAGCTGAGGCAAGAGATCGTTGAGCCAGGAGGTG	459
OY	673	GAGGTTGCAGTGAAGTGAATGTGTCCTACTCACTCAACTTGGGAGACAGATGACATT	732
Dd	460	GAGGTTGCAGTGAAGTGAATGACACCATGTGCACTCAGCTGGGGACACGAGGCAAACT	519
OY	733	TGTGTCAAAAGAAAAAAAAAAAAACAAGTAAAAAGAAACAGGTGAAGTTAACTTTAAATA	792
Dd	520	CTGTCTCAA	579
OY	793	AACCAATGATGCCAAATACATCATCTTTCAA	824
Dd	580	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	611

RESULT 9	
CR749224	
LOCUS	2449 bp mRNA linear HTC 19-AUG-2004
DEFINITION	Homo sapiens mRNA; CDNA DKFZp781C2298 (From clone DKFZp781C2298).
ACCESSION	CR749224

ACCESSION	CR749224	GI:51476179
VERSION	CR749224.1	
KEYWORDS	HTC.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
Eulaziora; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 2449)
Ostenwelder, B., Obermaier, B., Deutscherpau, S., Schallp, A.,
Mueser, H. W., Wäit, R., and C. Osterer, A. Pöb, C. Has, M. and

CONSRM TITLE	The German cDNA Consortium Direct Submission
JOURNAL	Submitted (17-AUG-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neubiberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center, Heidelberg, Germany, 69006-8800, Heidelberg, Germany

Research Center (DKFZ), Email s.wiemand@dkfz-hidelberg.de; sequenced by Medigenomix (Marlinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp781C2298) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: <http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneId=DKFZp781C2298> Further information about the clone and the sequencing project is available at <http://mips.gsf.de/projects/cdna/>.

FEATURES	location/Qualifiers
source	1. .249

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP781C2298"
/tissue_type="colon carcinoma"
/clone_lib="781 (synonym: hicc4) . Vector pSport1_Sfi; host: DH10B; sites SfiI + SfiIb"
/dev_stage="adult"
/note="hypothetical protein"
1..2449
gene

```

CDS /gene="DKFZp781C2298"
161. .862

```

/gene="DKFZp781C2298"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH18080.1"

```

```

/db xref="GI:5147610"
/translation="MLREEATKKSKEPEGMALPOGRILTFRDVAIEPFLSEBWKCLNIA
ORALRYALWLENTYRNTELEFVNDISSCKMCKEASAOGNETIHHGTLORHSHNLRDPC
FOELKRDYHNPEFMOEBRENGDEAPMTEIKELISGPNDRDHOAGNCKPIKIDGSGSF
HSHLPFLAHTPOPEAKIGNOVEXSIINASLILTSORISCSFKTINSNYGNNSLHSLSP
IÖKKGSTHERKTFPM"

```

Query Match	10.8%	Score 215.6	DB 3	Length 2449
Best Local Similarity	79.7%	Pred. No. 4	4e-24	
Matches 279	Conservative	0	Mismatches 69	Indels 2
				Gaps 2

428 TTAGTAGCCACATTAAAACAGGTAAAAAAGCTGGGCGCAGTGGCTCACCTGTAATCC 487

Db 1961 TTAGTGTTCCTGAAAAATGACTTAAAGCGCTGGGCGGCTACGCCCTGTATCC 2020

488 CAGCAGCTTTGGGAGGCTGAGGAGGAGAGATCACCTTTGGTACAGGAGTTTGAGACTAGACCT 547

Db 2021 CAGCACCTTTGGGAGGCTGAGGTAGTGGATCACCTGAGGTCAGGAGTTTGAGACCAAGCCT 2080

548 GGCACACATGGCGAACTCTGTCTTAAAAAATACAAAAATTAGCCTGGCATGGTGG 607

Db 2081 GACCAAAAAGTCAACTCCGTCTT-CTAAAAATACAAAAATTAGCCAGCATGTAG 2139

608 CCGGCGCCTGTATCTCAGCTGCTCAGAGGCCGAGACACAGAATCACTTAAACCAGG 667

Db 2140 CAGACGCGCTGTAGTTCACGCTATTCCAGGAGATGAGACAGAGAAATTCCTTGAACCCAGG 2199

668 AGGTGAGGTTGCAGTGAGCTGAGATCGTCCACT - CACTTCAACCTGGGAGACAGAGTG 726

Db 2200 AGGTGAGGCTGCAGTAAGCCGAGATCTTGCCACTGCACCTTAGCCCTGGGCAACAGAGTG 2259

Qy 727 ACACCTTTGTCCTCAAAAAGAAAAAACAAGTAAAAAAACAACAGGT 776

Db 2260 AGACTCTTGTCTCCAAAAAAGAAAGAAAAAGACAAAAAGAAACATGT 2309

RESULT 10

BX477553	573 bp	mRNA	linear	EST 04-SEP-2003
LOCUS				
BX477553				

DEFINITION
DKFZp686H1196_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone DKFZp686H1196_5', mRNA sequence.

ACCESSION	BX477553
VERSION	BX477553.1
	GI:31911608

KEYWORDS	EST.
SOURCE	Homo sapiens (human)

ORGANISM	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

REFERENCE 1 (bases 1 to 573)
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS
Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobó, G., Han, M. and Wiemann, S.

TITLE
EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)
JOURNAL
Unpublished (2003)

COMMENT Contact: MIPS
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by AGORA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.

No s1 sequence available.
This clone (DKFZp686H1196) is available at the RZPD in Berlin.

please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES	Location/Qualifiers
source	1. .573

```
/organism="Homo sapiens"  
/mol_type="mRNA"
```

```
/db_xref="taxon:9606"  
/clone="DKFZp686H11.96"
```

```
/dev_stage="adult"
```

/lab host="DH10B"
/clone_lib="686 (synonym: h1cc3)"
/note="Vector: pTriplEx2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN

Query Match 10.8%; Score 215.4; DB 5; Length 573;
Best Local Similarity 82.0%; Pred. No. 7.3e-24;
Matches 260; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

QY 464 CGCAGGCTTCACACCTGTAATCCAGACCTTGGAGGCTGAGCAGCAGATCACCTT 523
DB 1 CGCAGGCTTCACACCTGTAATCCAGACCTTGGAGGCTGAGCAGCAGATCACCTG 60
QY 524 TGGTCAGAGCTTGAAGCTAGCTGGCCCAATGGCCAAATCTGTCTCTTAAAAAAT 583
DB 61 AGGTCAAGAGTTCAAGACTAGCTGGCCCAATGGCCAAATCTGTCTCTTAAAAA 120
QY 584 ACAAAATTAGCTGCGATGGTGGCGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAG 643
DB 121 AAAAAATGAGCCAGCATGTGGCGGCACTGTAATCCAGCCGCTCGGAGGCTGAA 180
QY 644 ACACAGAAATCACTTAAACCCAGAGGTGAGGTTGAGTGAAGTGTGCTCACT- 702
DB 181 GCAGAGAAATCGCTTGAACCCAGAGGAGGTTGAGTGAAGTGTGCTGAGATGCACACTG 240
QY 703 CACTCCAACTGGGAGACAGATGACCTTTGTCTCAAAAAAAGAAAAAACAATA 762
DB 241 CACTCCAGCTGGGCGCAGATGACCTGTCTCAAAAAAAGAAAAAAGAAAAAAGA 300
QY 763 AAAAAAGAACAGGTGAA 779
DB 301 GAAAAACAGGAGAAAGA 317

RESULT 11 823 bp mRNA linear EST 15-MAY-2003
CD101954/c
LOCUS
DEFINITION
AGENCOURT 13980468 NIH_MGC_187 Homo sapiens cDNA clone
IMAGE:30373404 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 823)
NIH-MGC http://mgs.nci.nih.gov/.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgsaps-j@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDCML67 row: c column: 13
High quality sequence stop: 327.
Location/Qualifiers

FEATURES
source
1..823
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30373404"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_187"
/note="Organ: Blood vessels - aorta, basilar and artery;
Vector: pDNR-LIB; Site 1: SfiI (ggccatggcc); Site 2:
SfiI (ggccgctcgcc); 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence:
5'-CAGGCGCATATAGGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGGCGCCGACATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.4 kb
(range 0.5-4.0 kb). 14/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 10.8%; Score 215.4; DB 6; Length 823;
Best Local Similarity 80.8%; Pred. No. 6.5e-24;
Matches 263; Conservative 0; Mismatches 61; Indels 1; Gaps 1;

QY 444 AACAGTAAAAAAGGCTGGCGCAGTGGCTCAACCTGTATATCCAGACTTTGGAGGC 503
DB 326 AAGAAATTAATGAGGCGCAGGATGAGTGGCTCAACCTGTATATCCAGACTTTGGAGGC 267
QY 504 TGAAGCAGAGATCACTTGGTCAAGAGTTGAGACTAGCTGAGCCCAATGCGGAA 563
DB 266 TGAGGCGAGGATCACTTGAAGTCAAGATTGAAACAGCTTGGCCACAGCAAAA 207
QY 564 CTGTGCTTAAAAAATAAATAAATTAGCTGCGATGTGCGGCGCTGTATCT 623
DB 206 CCGTGTCTT-ACATAAATAAATAAATTAGCTGTATGTGTGCGGCGCTGTATCT 148
QY 624 CAGCTGCTCAGAGGCGGAGACCAAGATCACTTAAACCCAGAGGTGAGGTTGCACT 683
DB 147 CTGTACTTGGGAGGCTGAGACAGAAATCACTTGAACCCAGAGGTTAGGTTGCACT 88
QY 684 GAGCTGAGATCGGCGCACTCACTCCAACTGGGAGCAGAGTGAACCTTTGTCTCAAA 743
DB 87 GAGCTGAGATCAATCCACTGCACTAGCTGGGCAACAGCAAACTGTCTCAAAA 28
QY 744 AGAAAAAATAAATAAATAAATAAAG 768
DB 27 AAAAAAAAAAAAAAAAAAATG 3

RESULT 12 397 bp mRNA linear EST 19-OCT-2000
AV759632
LOCUS
DEFINITION
AV759632 MDS Homo sapiens cDNA clone MDSDBF02 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
EST.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 397)
AV759632
NIH-MGC http://mgs.nci.nih.gov/.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES
source
1..397
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MDSDBF02"
/tissue_type="Bone marrow"
/cell_type="CD34+ hematopoietic stem/progenitor cell"

```

/lab_host="BM25.8"
/clone_lib="WDS"
/note="Vector: pTriplex2, Site_1: sflra, Site_2: sflrb

```

Query Match	Similarity	Score	DB	Length
Best Local	80.1%	Pred. No. 9.4e-24;		
Matches	265;	Conservative	0;	Mismatches 65; Indels 1; Gaps 1

QY	458	GCTGGGCGCAGTGGCTCA	CACCTGTAA	TCCAGCA	CTTTGGAGGCTGAGCGACAGAT	517	
Db	1	GCCGGTCGACAGTGCTCA <td>CGCTGTAA <td>TCCAGCA <td>TTTGGGAGGCCGAGCGCAGGTGAT</td> <td>60</td> </td></td>	CGCTGTAA <td>TCCAGCA <td>TTTGGGAGGCCGAGCGCAGGTGAT</td> <td>60</td> </td>	TCCAGCA <td>TTTGGGAGGCCGAGCGCAGGTGAT</td> <td>60</td>	TTTGGGAGGCCGAGCGCAGGTGAT	60	
QY	518	CACCTTTGTCAGGAGTTTGAGACT <td>TAGCTGGCCAA</td> <td>CATGCGGAA <td>ACTTGTCTTAA</td> <td>577</td> </td>	TAGCTGGCCAA	CATGCGGAA <td>ACTTGTCTTAA</td> <td>577</td>	ACTTGTCTTAA	577	
Db	61	CACCTGAGGTGAGGAGTTTGAGAC	CGAGCTGGCCAA	CATGTGGAA	ACCCCGCTCTACTA	120	
QY	578	AAAAATACAAAAA	TTTAGCTGGCAT	GTGGCGGGCGCTGTAA	TCACAGTCTCAGAG	637	
Db	121	AAAAATACAAAAA	TTTAGCTGGCAT	GTGGCGGGCGCTGTAA	TCACAGTCTCAGAG	180	
QY	638	GCCGAGACACA	GAAATCA	CTTAAAC	CCGAGAGTGGAGTTGCA	TGAGCTGAGATCGTG	697
Db	181	GCTGAGGACAGAGAT	TCGCTTGA	ACTTGGGAGTGTGTGTTG	CACTGAGCCAA	GATCA	240
QY	698	CCACT	CACCTCCA	CTTGGAGACAGAGT	GACATTTTGTCTCA	AAAAAGAAAAAAA	756
Db	241	CCATTGCACTCCAGCTG	GTGGTGA	CGAGATGAGACTGTCT	TCAGAAAGAAAAAAA	300	
QY	757	CACGTAAAAA	AGAACAGGTGA	AGTTAACTT			
Db	301	AAAAAAATCGCGCCGCA	ATTTGAGTCTCACTT				

RESULT 13	BX644875/c	587 bp	mRNA	linear	EST 04-SEP-2001
LOCUS					
DEFINITION	DK22P781D1236.r1.761 (synonym: h1icc4) Homo sapiens cDNA clone				
ACCESSION	DK22P781D1236.5', mRNA sequence.				
VERSION	BX644875				
KEYWORDS	BX644875.1 GI:34479208				
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
REFERENCE	1 (bases 1 to 587)				
AUTHORS	Wambutt, R., Heubner, D., Mewes, H.W., weil, B., Amlid, C., Osanger, A.,				
	Pobo, G., Han, M., and Wiemann, S.				
TITLE	EST (Wambutt, R., Heubner, D., Mewes, H.W., weil, B., Amlid, C., et al.)				
JOURNAL	Unpublished (2003)				
COMMENT	Contact: MIPS				

FEATURES
source

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZP781D1236"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="781 (synonym: hicc4)"
note=vector: pSPORT1_Sfi, Site_1: SfiIA, site_2: SfiIB

```

cdna-collection™

Query Match	10.7%;	Score 215;	DB 5;	Length 587;
Best Local Similarity	80.1%;	Pred. No. 8,38-24;		
Matches	265;	Conservative	0;	Mismatches 65; Indels 1; Gaps 1;
QY	442	AAAACAGGTAAAAAGGCTGGGCGCAGTGGCTCAACCTGTAAATCCACAGCATTTTGGAG	501	
Db	463	AGAAAAAGAAAGAGAGCTGGGTGTGGTGCTCACGCTGTAAATCCAGTACTTTTGAAG	404	
QY	502	GCTGAGGCAGGCAACATCACTTTTGTGTCAGAGTTTGAACCTAGCTTGGCCAAATGCGGA	561	
Db	403	GCCAGGGGGGTGATCACTTGAAGTCAGGAGTTTGAAGCCAGGCTGGCCAAATGTGGA	344	
QY	562	AAAGCTGCTCTAAAAAAAATAGCAAAATTAAGCGTGGCATGTGGCGGCGGCTGTAT	621	
Db	343	ATTCTCTCTTATCAAAAAAATACAAAAATTAGCCAGGATGTGTGCAAGATGCTGTAT	284	
QY	622	CTCAGCTGCTCAGAGGCGGAGACACAGATCACTTAAACCCAGAGGTGAGTTGCA	681	
Db	283	CCGAGCTAAGTGGGAGGGGTGAGGACAGAGAAATCTTGAAACCGGAGAGCGGAGTTGCA	224	
QY	682	GTGAGCTGAGATCGTCCACT-CAGCTCAACCTGGGAGACAGAGTGAACCTTTGTCTCA	740	
Db	223	GTGAGTTGAGATTGTGCTACTGTACTCCAGGCTTGGGCAACAGAGAGACCTTATCCCA	164	
QY	741	AAAAAGAAAAAAAACAAGTAAAAAGAA	771	
Db	163	AAAAAAAAAAAAAGAAAAAGAAAAAGAAA	133	

RESULT	14
LOCUS	CK905714/c
DEFINITION	633 bp mRNA linear EST 11-MAR-2004 ij76h09.x6 Human insulinoma Homo sapiens cDNA clone IMAGE:5636992 3', mRNA sequence.
ACCESSION	CK905714
VERSION	CK905714.1
KEYWORDS	GI:45367245
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 633) Melton,D., Meadows,A., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A., Schmitt,C., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cadenas,M., Gibbons,M., McCamn,R., Cole,R., Tsagareishvili,R., Williams,T., Jackson,Y. and Bowers,Y.
TITLE	WashU-Harvard Pancreas EST Project
JOURNAL	Unpublished (2000)
COMMENT	Other ESTs: ij76h09.y1

FEATURES

source

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:5636992"  
/tissue_type="Insulinoma"
```

ORIGIN

/lab host="DH10B (phage-resistant)"
/clone lib="Human insulinoma"
/note="Organ: pancreas; Vector: plasmid SK-; Site: 1:
XhoI; Site 2: EcoRI; Constructed with lambda ZapII system
(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
plasmid SK- by Dr. H. Inoue following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Hiroshi Inoue, MD/PhD for further
information on this library (Metabolism Division, Permutt
Laboratory, Washington University School of Medicine, Box
8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this
is a Washington University Pancreas EST project library."

ORIGIN

Location/Qualifiers
1. 491
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2790849"
/issue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
/lab host="DH10B"
/clone lib="NCI CGAP_Ut2"
/note="Organ: uterus; Vector: pCMV-Sport6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: 0150 dt.
Average insert size 1.85 Kb. Life Technologies catalog #:
11539-012"

Query Match 10.7%; Score 215; DB 7; Length 633;
Best Local Similarity 83.4%; Pred. No. 8.1e-24;
Matches 256; Conservative 0; Mismatches 50; Indels 1; Gaps 1;

Query Match 10.7%; Score 214; DB 2; Length 491;
Best Local Similarity 79.4%; Pred. No. 1.3e-23;
Matches 278; Conservative 0; Mismatches 70; Indels 2; Gaps 2;

454 AAAGCTGGGCGGCTGCTGACACCTGATCCAGACTTTGGAGGCTGAGGACGC 513
308 ACAGGCGGGGCGGCTGCTGACACCTGATCCAGACTTTGGAGGCGGAGGAGGT 249
514 AGATCAGCTTTGCTCAGAGGTTGAGACTGAGCTGCGCAATGAGCACTGCTCT 573
248 GGATCAGCTGAGCTCAGAGGTTGAGACGACGCTGCGCAATGAGCACTGCTCT 189
574 AAAAAAATTCAAAAATTTAGCTGCGAGTGGCGGCGCTGATTTCTGAGTCTCA 633
188 ACTAAAAAATTCAAAAATTTAGCTGCGAGTGGCGCTGATTTCTGAGTCTCA 129
634 GAGGCGGAGGAGCAAGATCTTAAACCGAGAGGAGGAGTTCAGAGTGCATGAT 693
128 GAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 69
694 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
68 CACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 9
753 AAAAAA 759
8 AAAAAA 2

Db 753 AAAAAA 759

Db 8 AAAAAA 2

RESULT 15
AWS16097/c 491 bp mRNA linear EST 03-MAR-2000

LOCUS
DEFINITION
x60105.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2790849 3'
similar to contains Alu repetitive element, mRNA sequence.

ACCESSION
AWS16097
VERSION
AWS16097.1 GI:7154179
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
image.lnl.gov/image/ncml/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 424.

428 TTAGTACCCATTTAAACAGTAAAGGCTGGCGGCTGCTGCTGCTGCTGCTGCT 487
481 TTAGTACCCATTTAAACAGTAAAGGCTGGCGGCTGCTGCTGCTGCTGCTGCT 422
488 CAGACTTTGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 547
421 CAGACTTTGGAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 362
548 GGGCAGCATGCGCAAACTGCTGCTTAAAGGCTGGCGGCTGCTGCTGCTGCTG 607
361 GACCAAAAAGTCAAACTGCTGCTTAAAGGCTGGCGGCTGCTGCTGCTGCTG 303
608 GGGCGGCTGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 667
302 CAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
668 AGGAGGAGGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 726
242 AGGTGAGGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 183
727 ACACTTTGCTCAAAAGAAAAAACAATTAAGGCTGGCGGCTGCTGCTGCTGCT 776
182 AGACTTGTCTCAAAAGAAAAAACAATTAAGGCTGGCGGCTGCTGCTGCTGCT 133

Db 727 ACACTTTGCTCAAAAGAAAAAACAATTAAGGCTGGCGGCTGCTGCTGCTGCT 776

Db 182 AGACTTGTCTCAAAAGAAAAAACAATTAAGGCTGGCGGCTGCTGCTGCTGCT 133

Search completed: February 27, 2005, 12:42:21
Job time : 6279.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 ; Search time 795.397 Seconds
(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 aagagagatctgaggaatga.....ccagaataaactgagaag 1571

Scoring table: IDENTITY_NUC
gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Genesegq_16Dec04:*
2: genesegq1980s:*
3: genesegq1990s:*
4: genesegq2000s:*
5: genesegq2001as:*
6: genesegq2002as:*
7: genesegq2002bs:*
8: genesegq2003as:*
9: genesegq2003bs:*
10: genesegq2003cs:*
11: genesegq2003ds:*
12: genesegq2004as:*
13: genesegq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	11570	3	AAA95905 Human KLK
2	1571	100.0	11570	12	ADK52482 Human kal
3	1571	100.0	11570	13	ADR72623 Human ren
4	1571	100.0	11570	13	ADR72875 Human ova
5	541	34.4	567	12	ACH74409 Human gen
6	482	30.7	586	6	ABL63579 Breast ca
7	482	30.7	586	6	ABL63996 Breast ca
8	482	30.7	735	3	AAC79469 CDNA sequ
9	482	30.7	735	6	ABK29013 Human bre
10	482	30.7	738	2	AAK84240 DNA encod
11	482	30.7	1260	10	ADB80489 Ovarian c
12	482	30.7	1260	11	ADN39197 Cancer/an
13	482	30.7	1260	13	ADR25550
14	482	30.7	1302	2	AAK57990 Human BS2
15	482	30.7	1302	2	AAK57989 Human BS2
16	482	30.7	1381	2	AAZ06259 Human sec
17	482	30.7	1381	8	ADA40370 Human sec
18	482	30.7	1387	10	ADA56535 Gene enco
19	482	30.7	1387	13	ADR72622 Human ren
20	482	30.7	1387	13	ADR72874 Human ova

21	482	30.7	1439	2	AAZ06260 Human sec
22	482	30.7	1439	8	ADA40371 Human sec
23	482	30.7	1439	10	ADA56536 Gene enco
24	482	30.7	1441	8	ADA40368 Human sec
25	482	30.7	1441	10	ADA56533 Gene enco
26	482	30.7	1481	12	ADQ22212 Human sec
27	482	30.7	1490	8	ADA40367 Human sec
28	482	30.7	1490	10	ADA56532 Gene enco
29	482	30.7	1499	2	AAK60578 Human ker
30	482	30.7	1516	2	AAZ06245 Human sec
31	482	30.7	1516	2	ADA40369 Human sec
32	482	30.7	1516	10	ADA56534 Gene enco
33	482	30.7	1570	3	AAZ65070 Membrane-
34	482	30.7	1570	4	AAK521471 Human CDN
35	482	30.7	1570	5	AAK44216 Human PRO
36	482	30.7	1570	6	ABK28605 Human DNA
37	482	30.7	1570	8	ACA64363 Novel hum
38	482	30.7	1570	8	ACA03830 CDNA enco
39	482	30.7	1570	8	ABK89368 DNA enco
40	482	30.7	1570	8	ABK80822 Human sec
41	482	30.7	1570	8	ACD44331 CDNA enco
42	482	30.7	1570	8	ACD42022 Human sec
43	482	30.7	1570	8	ABK79502 Human sec
44	482	30.7	1570	8	ACA93523 Novel hum
45	482	30.7	1570	8	ABK81205 Human sec

ALIGNMENTS

RESULT 1

AAA95905
ID AAA95905 standard; DNA; 11570 BP.

XX AAA95905;

DT 02-FEB-2001 (first entry)

XX Human KLK-L2 gene.

XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;

KW kallikrein-like protein; serine protease; cytosolic; cancer;

XX prostrate cancer; ds.

XX Homo sapiens.

XX MO200053776-A2.

PD 14-SEP-2000.

PF 09-MAR-2000; 2000MO-CA000258.

PR 11-MAR-1999; 99US-0124260P.

PR 01-APR-1999; 99US-0127386P.

PR 21-JUL-1999; 99US-0144919P.

XX (MOUN) MOUNT SINAI HOSPITAL.

PI Yousef GM, Diamandis EP;

DR WPI; 2000-587440/55.

PT P-PSDB; AAB21296.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L

XX protein mediated disorders, especially cancer.

PS Claim 1; Page 143-149; 184pp; English.

The present sequence is the coding sequence of the human KLK-L2 gene, which encodes a kallikrein-like protein. Kallikreins and kallikrein-like proteins are a subgroup of the serine protease enzyme family. They catalyze the selective cleavage of specific polypeptide precursors to release peptides with potent biological activity. Nucleic acids encoding

CC kallikrein-like proteins KLK-11, KLK-12, KLK-13, KLK-14, KLK-15 and KLK-16 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1571; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No.0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AGGAGGTATGGGGAATTTGAAGACAGAAAACAAATTAAGTCCAGCGAATGGATTCTTA 60
DB 10000 AGGAGGTATGGGGAATTTGAAGACAGAAAACAAATTAAGTCCAGCGAATGGATTCTTA 10059
QY 61 TTGGAGGTATTTCTGCCCTAGAAACACTGGCAATACAGAGACATTTTGTGTCA 120
DB 10060 TTGGAGGTATTTCTGCCCTAGAAACACTGGCAATACAGAGACATTTTGTGTCA 10119
QY 121 CAACATATGAGGGGCAATTACTGAGCACTAATGATAGATGCAAGTGTCTGTCAAC 180
DB 10120 CAACATATGAGGGGCAATTACTGAGCACTAATGATAGATGCAAGTGTCTGTCAAC 10179
QY 181 ATGCTATGATGACACAGGAGGCTCCACAAACCAATTATCCAGTTCAGATGCCAC 240
DB 10180 ATGCTATGATGACACAGGAGGCTCCACAAACCAATTATCCAGTTCAGATGCCAC 10239
QY 241 AGTGCCCAATCGAGGAACCTTCATCGAGGGCTGAAACCGTATTTTTCAGAGAGAG 300
DB 10240 AGTGCCCAATCGAGGAACCTTCATCGAGGGCTGAAACCGTATTTTTCAGAGAGAG 10299
QY 301 GATTAAGATGGGTGTGGAGAAATGGGGAAGGAGTGTGTCCAGTAAGAAATAA 360
DB 10300 GATTAAGATGGGTGTGGAGAAATGGGGAAGGAGTGTGTCCAGTAAGAAATAA 10359
QY 361 GGGCTTGACAGAGCTGAGGGGAGAGTGAAGAGAAAAGGAGCGAGATACACAGATGA 420
DB 10360 GGGCTTGACAGAGCTGAGGGGAGAGTGAAGAGAAAAGGAGCGAGATACACAGATGA 10419
QY 421 GGGAGACAGGCTGGAACAGAAAGTGAAGACGAAATTCAGATGTGAAGAGAGGCTCA 480
DB 10420 GGGAGACAGGCTGGAACAGAAAGTGAAGACGAAATTCAGATGTGAAGAGAGGCTCA 10479
QY 481 CAGACCCCCCAAAATGATGTGACAAACAGGAATCTGGAAGAGAAAGATGGAAGTGGAG 540
DB 10480 CAGACCCCCCAAAATGATGTGACAAACAGGAATCTGGAAGAGAAAGATGGAAGTGGAG 10539
QY 541 AGTGACAAATGGGCTTAAAGGTTGAACCTTGAGGCGCAGGCAATGTGCTCAAGCTCTTA 600
DB 10540 AGTGACAAATGGGCTTAAAGGTTGAACCTTGAGGCGCAGGCAATGTGCTCAAGCTCTTA 10599
QY 601 ATCCCAACACTTTGAGGCTGAGGTGGGGAATCACTTGAAGCCAGAGATTGAGACAG 660
DB 10600 ATCCCAACACTTTGAGGCTGAGGTGGGGAATCACTTGAAGCCAGAGATTGAGACAG 10659
QY 661 CCGGCGCAACATGGTGAACCCCGTCTTACAAAATAAATTAACAAATTAAGCCGGGTGT 720
DB 10660 CCGGCGCAACATGGTGAACCCCGTCTTACAAAATAAATTAACAAATTAAGCCGGGTGT 10719
QY 721 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGCGAGAGATTGCTTGAAC 780
DB 10720 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGCGAGAGATTGCTTGAAC 10779
QY 781 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCTCACTGCGCTCAACTGGGCAACA 840
DB 10780 CCGGAGATGAGAGGCTGAGTGAAGTGAAGTCAAGGCTCACTGCGCTCAACTGGGCAACA 10839
QY 841 GAGTAAAGCTCAATCTCAAAAAAATAAAGCTGGAATTTGAAGTAATTAATTAACATT 900
DB 10840 GAGTAAAGCTCAATCTCAAAAAAATAAAGCTGGAATTTGAAGTAATTAATTAACATT 10899
```

```
QY 901 CTCCTCTCTCTCCTTTTGGCTGTGTCTTCATCTGTCTTTTTCGACATTTCTTCATCT 960
DB 10900 CTCCTCTCTCTCCTTTTGGCTGTGTCTTCATCTGTCTTTTTCGACATTTCTTCATCT 10959
QY 961 CTGTACTTTCATCTCTGTGTGTCTGTTCCTCATCTGTCTTTCATCTATAGGCAATCTCG 1020
DB 10960 CTGTACTTTCATCTCTGTGTGTCTGTTCCTCATCTGTCTTTCATCTATAGGCAATCTCG 11019
QY 1021 GGTCTCTATGTCTCTCTTGGCCCACTTGGCAATCTGTGCTCTCATGATCCCCCTT 1080
DB 11020 GGTCTCTATGTCTCTCTTGGCCCACTTGGCAATCTGTGCTCTCATGATCCCCCTT 11079
QY 1081 TCTCTCTGACAGGATGATTTCTGGGGGCGCTGTGTCTGCAATGCTCTCCGACGAGACTC 1140
DB 11080 TCTCTCTGACAGGATGATTTCTGGGGGCGCTGTGTCTGCAATGCTCTCCGACGAGACTC 11139
QY 1141 GTGTCTGGGAGATTTACCTTGTGTGCTGGGCCCAACAGACGGGAGTCTACACGAACCTC 1200
DB 11140 GTGTCTGGGAGATTTACCTTGTGTGCTGGGCCCAACAGACGGGAGTCTACACGAACCTC 11199
QY 1201 TGCAGTTTCAACCAAGTGAATCCAGGAAACATCCAGGCGCACTCTGAGTCAATCCAGGA 1260
DB 11200 TGCAGTTTCAACCAAGTGAATCCAGGAAACATCCAGGCGCACTCTGAGTCAATCCAGGA 11259
QY 1261 CTCAGCACACCGGCATCCCACTGCTGACAGGACAGGCTTGACACTCTCTTTCAGACCT 1320
DB 11260 CTCAGCACACCGGCATCCCACTGCTGACAGGACAGGCTTGACACTCTCTTTCAGACCT 11319
QY 1321 CATTCCTTCCAGAGATGTTGAATGTTTCACTCTTCAGAGCCCTGACCCCATGTCTCT 1380
DB 11320 CATTCCTTCCAGAGATGTTGAATGTTTCACTCTTCAGAGCCCTGACCCCATGTCTCTCT 11379
QY 1381 GGACTCAGGGGTGTGCTTCCCCCAATGAGGCTGACAGTGTCTCTAGTTGAACCTGGG 1440
DB 11380 GGACTCAGGGGTGTGCTTCCCCCAATGAGGCTGACAGTGTCTCTAGTTGAACCTGGG 11439
QY 1441 AACCAATTTCCAAACTGTCCAGGGGCGGGGTTGCGTCAATCTCTCGGGCACTTTCA 1500
DB 11440 AACCAATTTCCAAACTGTCCAGGGGCGGGGTTGCGTCAATCTCTCGGGCACTTTCA 11499
QY 1501 TCTCTAAGCTCAGGGCCCATCTCTTCTGTGACGCTGTGACCCAAATTTAGTCCGAAAT 1560
DB 11500 TCTCTAAGCTCAGGGCCCATCTCTTCTGTGACGCTGTGACCCAAATTTAGTCCGAAAT 11559
QY 1561 AAACTGAGAAG 1571
DB 11560 AAACTGAGAAG 11570
```

RESULT 2
ADKS2482
ID ADKS2482 strand: DNA; 11570 BP.
XX
AC ADKS2482;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human kallikrein 5 encoding sequence.
XX
KW kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.
XX
XX Homo sapiens.
OS
PN WO2004021008-A2.
XX
PD 11-MAR-2004.
XX
PF 28-AUG-2003; 2003WO-CA001310.
XX
PR 28-AUG-2002; 2002US-0407333P.
XX
XX (MOUN) MOUNT SINAI HOSPITAL.
PA (YOUS /) YOUSEF G.

XX Diamond E³;
PI WPI; 2004-239232/22.
DR P-PSDB; ADK52481.
XX
PT Detecting kallikrein 5 associated with cancer, useful in diagnosing,
PT monitoring, detecting, imaging and treating breast or ovarian carcinoma,
PT comprises comparing the detected amount of kallikrein 5 in a sample with
PT a standard sample.
XX
PS Disclosure; SEQ ID NO 2; 59pd; English.
XX
CC The present invention relates to detecting kallikrein 5 associated with
CC breast or ovarian cancer in a patient comprises detecting in the sample
CC kallikrein 5 and comparing the detected amount with an amount detected
CC for a standard. The method is useful in detecting kallikrein 5 associated
CC with breast or ovarian cancer in a patient. The methods and kits are
CC useful in diagnosing, monitoring, detecting, imaging and treating breast
CC or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for
CC preventing and treating breast and ovarian cancer and for stimulating or
CC enhancing antibody production or for inducing an immune response. The
CC present sequence represents human kallikrein 5 encoding sequence.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 1571; DB 12; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAGGTATGGGAGTTGAAGACAGAAACAAATTAATCCAGGAAATGATTTCTA 60
DB 10000 AGGAGGTATGGGAGTTGAAGACAGAAACAAATTAATCCAGGAAATGATTTCTA 10059
QY 61 TTGGAGTATGTCGCCCTAGAAACACTGCAATCCAGAGACATTTTGGTGTCA 120
DB 10060 TTGGAGTATGTCGCCCTAGAAACACTGCGCATCCAGAGACATTTTGGTGTCA 10119
QY 121 CAATATATGAGGGGCAATTAATGATGATGATGCCAAGTGTGTTCAAC 180
DB 10120 CAATATATGAGGGGCAATTAATGATGATGATGCCAAGTGTGTTCAAC 10179
QY 181 ATCTATGATGACACGGGAGGCTCCCAACAAACATTAATCCAGGAAATGATTTCTA 240
DB 10180 ATCTATGATGACACGGGAGGCTCCCAACAAACATTAATCCAGGAAATGATTTCTA 10239
QY 241 AGTCCAGATTCAGAGAACCTCATCCAGGGGCTGAGAACGTAATTTTTCAGAAAGGAG 300
DB 10240 AGTCCAGATTCAGAGAACCTCATCCAGGGGCTGAGAACGTAATTTTTCAGAAAGGAG 10299
QY 301 GTATTAAGATGGGTGTGAGAAATGGGAAAGAGGTGTGTCCAGTAAGAAATAA 360
DB 10300 GTATTAAGATGGGTGTGAGAAATGGGAAAGAGGTGTGTCCAGTAAGAAATAA 10359
QY 361 GGGCTGACAGGCTGAGGGGAGAGTGAAGAAAGGAGGGGAGAGATTAACAGATGA 420
DB 10360 GGGCTGACAGGCTGAGGGGAGAGTGAAGAAAGGAGGGGAGAGATTAACAGATGA 10419
QY 421 GGGAGACAGGCTGAGAACAGAAAGTGAAGAGAAATTCAGATGTGAGAGAGAGGTCA 480
DB 10420 GGGAGACAGGCTGAGAACAGAAAGTGAAGAGAAATTCAGATGTGAGAGAGAGGTCA 10479
QY 481 CAGACCCCCCAATGATGTGTGACACAGAAATTCGAAAGAGAAAGATGAGTGTAG 540
DB 10480 CAGACCCCCCAATGATGTGTGACACAGAAATTCGAAAGAGAAAGATGAGTGTAG 10539
QY 541 AGTGAACAATGGGTCTAAAGTTGAATCTGAGGCGCAGGCAATGTGGCTCAGCCCTGA 600
DB 10540 AGTGAACAATGGGTCTAAAGTTGAATCTGAGGCGCAGGCAATGTGGCTCAGCCCTGA 10599
QY 601 ATCCCAACATTTGAGAGCTGAGGTGGGCAATCACTTGAGGCGCAGGAGTTTCAGAGACAG 660
DB 10600 ATCCCAACATTTGAGAGCTGAGGTGGGCAATCACTTGAGGCGCAGGAGTTTCAGAGACAG 10659

QY 661 CTTGGCCAAATGCTGAAACCCCTCTCTACAAAAAATTAACAAAAATTAAGCCGGGTGT 720
DB 10660 CTTGGCCAAATGCTGAAACCCCTCTCTACAAAAAATTAACAAAAATTAAGCCGGGTGT 10719
QY 721 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGAGAGAAATTCCTTGAAC 780
DB 10720 GGTGATGACACCTGTAGTCAACAGCTACTTGGAGGCTGAGGAGAGAAATTCCTTGAAC 10779
QY 781 CCGGAGATGAGAGCTGACAGTGAAGTCAAGGCTCAAGGCTCAACCTGGGCAACA 840
DB 10780 CCGGAGATGAGAGCTGACAGTGAAGTCAAGGCTCAAGGCTCAACCTGGGCAACA 10839
QY 841 GAGTAAGATCCATCTCAAAAAAAGCTGGAATTTGAGAGAAATTAATTAATTAATTAAT 900
DB 10840 GAGTAAGATCCATCTCAAAAAAAGCTGGAATTTGAGAGAAATTAATTAATTAATTAAT 10899
QY 901 CTCCT 960
DB 10900 CTCCT 10959
QY 961 CTGTACTTTCATCTCTGTGTGTCTGTTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
DB 10960 CTGTACTTTCATCTCTGTGTGTCTGTTCATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11019
QY 1021 GGTCTCTCATCT 1080
DB 11020 GGTCTCTCATCT 11079
QY 1081 TCTCTCTCTCAAGGATATCTGGGGGGGCTGTGTGTCTGCAATGAGTCTCTCTCTCTCTCTCT 1140
DB 11080 TCTCTCTCTCAAGGATATCTGGGGGGGCTGTGTGTCTGCAATGAGTCTCTCTCTCTCTCTCT 11139
QY 1141 GTGTCTCTGGGAGATTAATCTGT 1200
DB 11140 GTGTCTCTGGGAGATTAATCTGT 11199
QY 1201 TGAATGTTCAACCAAGGATTCAGAGAAACATTCAGAGCCCACTCTCTGAGTCAATCCAGAG 1260
DB 11200 TGAATGTTCAACCAAGGATTCAGAGAAACATTCAGAGCCCACTCTCTGAGTCAATCCAGAG 11259
QY 1261 CTCAGACACCGGATCCCACTCTGTGAGAGGACAGCCCTGACATCTCTCTCTCTCTCTCTCT 1320
DB 11260 CTCAGACACCGGATCCCACTCTGTGAGAGGACAGCCCTGACATCTCTCTCTCTCTCTCTCT 11319
QY 1321 CATTCCTTCCAGAGATGTTGAGAAATGTTCAATCTCTCAAGCCCTGACATCTCTCTCTCT 1380
DB 11320 CATTCCTTCCAGAGATGTTGAGAAATGTTCAATCTCTCAAGCCCTGACATCTCTCTCTCTCT 11379
QY 1381 GGAATCAAGGATGCTTCT 1440
DB 11380 GGAATCAAGGATGCTTCT 11439
QY 1441 AACATTTTCAAAAACTGTCAAGGGGAGGTTGCTGCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1500
DB 11440 AACATTTTCAAAAACTGTCAAGGGGAGGTTGCTGCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 11499
QY 1501 TCCTCAAGCTCAAGGCGCAATCT 1560
DB 11500 TCCTCAAGCTCAAGGCGCAATCT 11559
QY 1561 AAACGTAGAG 1571
DB 11560 AAACGTAGAG 11570
RESULT 3
ADR72623 standard; DNA; 11570 BP.
XX
XX ADR72623;
DT 02-DEC-2004 (first entry)

XX		Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.
XX	KW	kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;
KM		cancer metastasis; chemotherapy; human; serine protease;
KX		chromosome 19q13.4; KLK5; ds; gene.
XX	OS	Homo sapiens.
FH	Key	Location/Qualifiers
FT	CDS	2221..11247
FT		/tag= b
FT		/product= "Human renal cell carcinoma-related kallikrein
FT		5 (klk5), protein"
FT	exon	2221..2293
FT		/tag= a
FT		/number= 1
FT	intron	2294..4761
FT		/tag= c
FT		/number= 1
FT	exon	4762..5023
FT		/tag= d
FT		/number= 2
FT	intron	5024..5762
FT		/tag= e
FT		/number= 2
FT	exon	5763..6019
FT		/tag= f
FT		/number= 3
FT	intron	6020..6104
FT		/tag= g
FT		/number= 3
FT	exon	6105..6238
FT		/tag= h
FT		/number= 4
FT	exon	6239..11091
FT		/tag= i
FT		/number= 4
FT	exon	11092..11247
FT		/tag= j
FT		/number= 5
PN		WO2004077060-A2.
PD		10-SEP-2004.
PF		26-FEB-2004; 2004MO-CA000280.
PR		27-FEB-2003; 2003US-0451382P.
PA	(MOUN)	MOUNT SINAI HOSPITAL.
PI	Diamonds EP,	Petraki CD;
DR	WPI; 2004-662077/64.	
DR	P-PsDB; ADR72621.	
XX		Detecting kallikrein polypeptides/polynucleotides associated with renal
PT		cell carcinoma in a patient, for diagnosing/treating the disease,
PT		comprises detecting /identifying kallikrein polypeptides/polynucleotides
PT		in a sample.
XX		
PS	Example 1;	SEQ ID NO 3; s3pp; English.
XX		The invention relates to a novel method for detecting kallikrein
CC		polypeptides, or the polynucleotides encoding them, associated with renal
CC		cell carcinoma. The method comprises obtaining a sample from a patient
CC		and detecting kallikrein polypeptides, or their encoding polynucleotides,
CC		where the kallikrein polypeptides are selected from the group consisting
CC		of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The
CC		detected amounts of the kallikrein polypeptides are compared to standard
CC		amounts. The molecules of the invention demonstrate cytosolic activity
CC		whilst the methods and kit may be useful for detecting, characterising,

CC preventing and treating renal cell carcinoma. Furthermore, the methods
CC may be useful for evaluating the probability of the presence of malignant
CC or pre-malignant cells and for detecting and quantitating tumour growth
CC and cancer metastasis. Finally, the methods may be utilised to confirm
CC the absence or removal of all tumour tissue following surgery, cancer
CC chemotherapy and/or radiation therapy and to monitor cancer chemotherapy
CC and tumour reappearance. The current sequence is that of the human
CC ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC invention which encodes a secreted serine protease and is located at
CC chromosome 19q13.4.

XX
SQ Sequence 11570 BP, 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 1571; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AGGAGGATATGGGAATTGGAGACAGAAACA	CAAATTAGTCAAGCGAATGGATTTCTA	60	
Db	1000	AGGAGGATATGGGAATTGGAGACAGAAACA	CAAATTAGTCAAGCGAATGGATTTCTA	10059	
QY	61	TTGGGAGTGAATTTCTGCCCTTAGAAGACA	CTGGCAATACCAGAGACATTTTGTGTGTCA	120	
Db	10060	TTGGGAGTGAATTTCTGCCCTTAGAAGACA	CTGGCAATACCAGAGACATTTTGTGTGTCA	10119	
QY	121	CAACTATATGAGAGGGGCATTACCTGGCA	CTAATGATAGATGTCGCAAGTGTGTTCAAC	180	
Db	10120	CAACTATATGAGAGGGGCATTACCTGGCA	CTAATGATAGATGTCGCAAGTGTGTTCAAC	10179	
QY	181	ATGCTATGATGCA	CACGGCAGGCTTCCACA	CAAAACATTAATTCAGTTCAGATGCCAC	240
Db	10180	ATGCTATGATGCA	CACGGCAGGCTTCCACA	CAAAACATTAATTCAGTTCAGATGCCAC	10239
QY	241	AGTGCCCAAGTATGAGGAACCTTCATC	CAGGGCTGGAACCGTATTTTTCAGAAAGGAG	300	
Db	10240	AGTGCCCAAGTATGAGGAACCTTCATC	CAGGGCTGGAACCGTATTTTTCAGAAAGGAG	10239	
QY	301	GTATTAAGATGGGTGTGTGTGAGAAATGGG	AGAAAGTGTGTGTCCAGTAAGAAATTA	360	
Db	10300	GTATTAAGATGGGTGTGTGTGAGAAATGGG	AGAAAGTGTGTGTCCAGTAAGAAATTA	10359	
QY	361	GGCCTGCA	CAGGCTGAGGGGAGAGTGAAGAGAAAGGA	GGCGGAGAGATTCACAGATGA	420
Db	10360	GGCCTGCA	CAGGCTGAGGGGAGAGTGAAGAGAAAGGA	GGCGGAGAGATTCACAGATGA	10419
QY	421	GGGAGACAGGCTGGAACAGAAAGTAGAC	AGAAAGTTCAGATGTGTGAGAGGAAGGCTCA	480	
Db	10420	GGGAGACAGGCTGGAACAGAAAGTAGAC	AGAAAGTTCAGATGTGTGAGAGGAAGGCTCA	10479	
QY	481	CAGACCCCCCGAAATGATGTGTGACA	CAGGAATCTGGAAAGAGAGATGAGTGAAG	540	
Db	10480	CAGACCCCCCGAAATGATGTGTGACA	CAGGAATCTGGAAAGAGAGATGAGTGAAG	10539	
QY	541	AGTGACAAATGGGGCTTAAAGGTGGA	CTTGAAGGCAAGGCANTGTGGCTCA	CAGGCTGTA	600
Db	10540	AGTGACAAATGGGGCTTAAAGGTGGA	CTTGAAGGCAAGGCANTGTGGCTCA	CAGGCTGTA	10539
QY	601	ATCCCAACATTTTGAGAGGCTGAGGTGG	CGAAATCACTTGAAGCCAGAGATTGACACGAG	660	
Db	10600	ATCCCAACATTTTGAGAGGCTGAGGTGG	CGAAATCACTTGAAGCCAGAGATTGACACGAG	10659	
QY	661	CTTGCCCAACATGTGTGAACCCCGTCT	CTACAAAAAAAATACAAAAAATTAGCCGGGTGT	720	
Db	10660	CTTGCCCAACATGTGTGAACCCCGTCT	CTACAAAAAAAATACAAAAAATTAGCCGGGTGT	10719	
QY	721	GGTGAATGGAACCTGTAGTCA	CAGGCTACTTGGGAGGCTGAGCAGAGAAATTGCTTGAAC	780	
Db	10720	GGTGAATGGAACCTGTAGTCA	CAGGCTACTTGGGAGGCTGAGCAGAGAAATTGCTTGAAC	10779	
QY	781	CCGGAGATGAGAGGCTGCAGTGAAGCTGA	GAGGCACTGGGCTTCAACTTGGGCAACA	840	
Db	10780	CCGGAGATGAGAGGCTGCAGTGAAGCTGA	GAGGCACTGGGCTTCAACTTGGGCAACA	10839	

Oy	841	GAGTAAAGACCTCAATCTGAAAAAATAAAGCTGGAATTGGAGGAATATTAATACATT	900
Db	10840	GAGTAAAGACCTCAATCTGAAAAAATAAAGCTGGAATTGGAGGAATATTAATACATT	108999
Oy	901	CTCCCTCTCTCTCCCTTTGGCCCTGTGTCTCATCTGTCTCTTTTCTGCATTTCTTCATCT	960
Db	10900	CTCCCTCTCTCTCTCTTTTGGCCCTGTGTCTCATCTGTCTCTTTTCTGCATTTCTTCATCT	109558
Oy	961	CTGTACTTTTCATCTGTGTGTGTCTGTTCCTCATCTGCTTCTCATCTAATGGGATCTGTG	1020
Db	10960	CTGTACTTTTCATCTGTGTGTGTGTTCCTCATCTGCTTCTCATCTAATGGGATCTGTG	11019
Oy	1021	GGTCTCTCAGTCTCTCTTCTTGCCCACTTTGCGACATCTCTGCTCTCATGCCCCCTT	1080
Db	11020	GGTCTCTCAGTCTCTCTTCTTGCCCACTTTGCGACATCTCTGCTCTCATGCCCCCTT	11079
Oy	1081	TCTCTCTGACGGGTGATTTCTGGGGGGGCGTGTGGTCTGCATAGGCTCCCTGACAGGACTC	1140
Db	11080	TCTCTCTCTGACGGGTGATTTCTGGGGGGGCGTGTGGTCTGCATAGGCTCCCTGACAGGACTC	11139
Oy	1141	GTGTCTCTGGGAGATTACCTTGTGTGCCGGGCCCAACAGACGGGGTGTCTACAGAACTCT	1200
Db	11140	GTGTCTCTGGGAGATTACCTTGTGTGCCGGGCCCAACAGACGGGGTGTCTACAGAACTCT	11139
Oy	1201	TGCAAGTTTCACCAAGTGGATTCACAGAAACCATTCAGGCCAACTCTCTGAATCTCCAGGA	1260
Db	11300	TGCAAGTTTCACCAAGTGGATTCACAGAAACCATTCAGGCCAACTCTGAATCTCCAGGA	11259
Oy	1261	CTGAGCACACCGGGATCCCAACCTGTGACAGGGAGACGCGCTGACATCTCTTTCAGACCT	1320
Db	11260	CTGAGCACACCGGGATCCCAACCTGTGACAGGGAGACGCGCTGACATCTCTTTCAGACCT	11319
Oy	1321	CATTCCTTCCAGAGATGTGAGATGTTTATCTCTTCAGCGCCCTGACCCCATGTCTCT	1380
Db	11320	CATTCCTTCCAGAGATGTGAGATGTTTATCTCTTCAGCGCCCTGACCCCATGTCTCT	11379
Oy	1381	GGACTCAGGGTCTCTCTTCCCCACATTTGGGGCTGACCGTGTCTCTAATTGAACCTTGGG	1440
Db	11380	GGACTCAGGGTCTCTCTTCCCCACATTTGGGGCTGACCGTGTCTCTAATTGAACCTTGGG	11439
Oy	1441	AACAAATTTCCAAAATCTGTCCAGGGCGGGGGTGGTCTCAATCTCCCTGGGGCATTTGA	1500
Db	11440	AACAAATTTCCAAAATCTGTCCAGGGCGGGGGTGGTCTCAATCTCCCTGGGGCATTTGA	11499
Oy	1501	TCCTCAAGCTCAGGGCCCATCTCTTCTCTGACGCTGCACCAATTTAGTCCCAAGAAAT	1560
Db	11500	TCCTCAAGCTCAGGGCCCATCTCTTCTCTGACGCTGCACCAATTTAGTCCCAAGAAAT	11559
Oy	1561	AAACTGAGAG 1571	
Db	11560	AAACTGAGAG 11570	
RESULT 4			
ADR72875			
AD ID	ADR72875	standard; DNA; 11570 BP.	
XX	AC	ADR72875;	
XX	DT	02-DEC-2004 (first entry)	
XX	DE	Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2.	
XX	KM	kallikrein 5; tumour marker; ovarian cancer;	
XX	KM	epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;	
XX	KLK5; ds; gene.		
OS	XX	Homo sapiens.	
PH	XX		
FT	Key	Location/Qualifiers	
FT	CDS	2221..11247	
FT	/*tag= b /product= "Human ovarian cancer-related tumour marker		

Query Match	Best Local Similarity	Matches	1 AGGAGGATTCGGGAAATTGAAGACAGAAAACACAAATTAAGTTCACAGCGAATGATTTCTTA	60
100.0%;	Score 1571; DB 13; Length 11570;			
100.0%;	Pred. No. 0;			
0;	Mismatches	0;	Indels	0; Gaps
0;				

DB 10000 AGGAGATATGAGGAATTGGAAGACAGAAACAAATATGATCCAAAGCAATGATTTCTTA 10059
QY 61 TTGGAGATGATTTGCCCCCTAGAAAGACATCGCAATACAGAGACATTTTGGTTGTA 120
DB 10060 TTGGAGATGATTTGCCCCCTAGAAAGACATCGCAATACAGAGACATTTTGGTTGTA 10119
QY 121 CAACATATATGAGAGGAGCATTTACTGAGCACTAATGATAGATGCAAGTGTCTGTTCAAC 180
DB 10120 CAACATATATGAGAGGAGCATTTACTGAGCACTAATGATAGATGCAAGTGTCTGTTCAAC 10179
QY 181 ATGCTATGATGCAACAGGAGGCTCCACAAACCAATATATGATGCAAGTGTCTGTTCAAC 240
DB 10180 ATGCTATGATGCAACAGGAGGCTCCACAAACCAATATATGATGCAAGTGTCTGTTCAAC 10239
QY 241 AGTGCACAGATGAGGACCCCTCATCGAGGAGCTGAAACCGATTTTGGCAAGAGGAG 300
DB 10240 AGTGCACAGATGAGGACCCCTCATCGAGGAGCTGAAACCGATTTTGGCAAGAGGAG 10299
QY 301 GTATAGATGATGAGTGTGAGAAATGAGGAGAGAGTGTGTCTCAGTAAAGAAATTA 360
DB 10300 GTATAGATGATGAGTGTGAGAAATGAGGAGAGAGTGTGTCTCAGTAAAGAAATTA 10359
QY 361 GGCCTGCAAGGCTGAGAGGAGAGATGAGAGAAAGGAGGCGAGAGATACAGATGA 420
DB 10360 GGCCTGCAAGGCTGAGAGGAGAGATGAGAGAAAGGAGGCGAGAGATACAGATGA 10419
QY 421 GGGAGACAGGCTGAGAACAGAAATAGAGAGAAATTCAGATGTGAGAGAGGAGTGA 480
DB 10420 GGGAGACAGGCTGAGAACAGAAATAGAGAGAAATTCAGATGTGAGAGAGGAGTGA 10479
QY 481 CAGACCCCCCGAAATGATGTGTGACAAACAGAAATCTGAGAGAGAGATGAGAGTGA 540
DB 10480 CAGACCCCCCGAAATGATGTGTGACAAACAGAAATCTGAGAGAGAGATGAGAGTGA 10539
QY 541 AGTGAACAATGGGCTCTAAAGTTGAACCTTGAGGCGCAGGATGTGGCTCACGCTGTA 600
DB 10540 AGTGAACAATGGGCTCTAAAGTTGAACCTTGAGGCGCAGGATGTGGCTCACGCTGTA 10599
QY 601 ATCCCAACCTTTGAGAGGCTGAGAGGAGGAGAAATCAGTTGAGCAGAGAGTTCAGACAG 660
DB 10600 ATCCCAACCTTTGAGAGGCTGAGAGGAGGAGAAATCAGTTGAGCAGAGAGTTCAGACAG 10659
QY 661 CCTGGCCAAATGATGTAACCCCGTCTTACAAAAAAAATACAAAAAATTAGCGGGGTGT 720
DB 10660 CCTGGCCAAATGATGTAACCCCGTCTTACAAAAAAAATACAAAAAATTAGCGGGGTGT 10719
QY 721 GGTGATGAGACCTGTATGTACAGCTAATTGGAGGCTGAGGAGAGAAATTCGTTGAAC 780
DB 10720 GGTGATGAGACCTGTATGTACAGCTAATTGGAGGCTGAGGAGAGAAATTCGTTGAAC 10779
QY 781 CCGGAGATGAGAGGCTGAGAGTGAAGTCAAGGCACTGCGCTCAACCTGGGAGACA 840
DB 10780 CCGGAGATGAGAGGCTGAGAGTGAAGTCAAGGCACTGCGCTCAACCTGGGAGACA 10839
QY 841 GAGTAAGACTCACTCAAAAAAAGCTGATTTGAGAGTGAATAATTAATTAACATT 900
DB 10840 GAGTAAGACTCACTCAAAAAAAGCTGATTTGAGAGTGAATAATTAATTAACATT 10899
QY 901 CTGCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGCAATTTCTTCATCT 960
DB 10900 CTGCTCTCTCTCTTTTGGCTGTCTCATCTGTCTTTTTCGCAATTTCTTCATCT 10959
QY 961 CTGTACTTTCCATCTGATGTGTCTGTTCCCATCTGCTTCATCTAATGAGGAGATCTG 1020
DB 10960 CTGTACTTTCCATCTGATGTGTCTGTTCCCATCTGCTTCATCTAATGAGGAGATCTG 11019
QY 1021 GGTCTCTCAATGCTCTCTTCTGCCCCCATTTGCAATCTGTGCTCTCAATGCCCCCTT 1080
DB 11020 GGTCTCTCAATGCTCTCTTCTGCCCCCATTTGCAATCTGTGCTCTCAATGCCCCCTT 11079
QY 1081 TCTCTCTGCAAGGATGATTTGAGGAGGCTGTGGTCTGCAATGGCTCCCTGCAAGGAGCTC 1140

DB 11080 TCTCTCTGCAAGGATGATTTGAGGAGGCTGTGGTCTGCAATGGCTCCCTGCAAGGAGCTC 11139
QY 1141 GGTCTCTGAGGAGATTAACCTTGTGCTGCGGCCCAACAGACCGGGTGTCTACAGAACTTC 1200
DB 11140 GGTCTCTGAGGAGATTAACCTTGTGCTGCGGCCCAACAGACCGGGTGTCTACAGAACTTC 11199
QY 1201 TGCAGATTCACCAAGTGAATCCAGAAACCAATCCAGGCCCACTCTGATCATCCAGGA 1260
DB 11200 TGCAGATTCACCAAGTGAATCCAGAAACCAATCCAGGCCCACTCTGATCATCCAGGA 11259
QY 1261 CTCAGACACCGGAGATCCCACTGCTGAGGAGACAGCCCTGACACTCTTTCAACCTT 1320
DB 11260 CTCAGACACCGGAGATCCCACTGCTGAGGAGACAGCCCTGACACTCTTTCAACCTT 11319
QY 1321 CATTCCTTCCAGAGATGTTGAAATGTTCACTCTCCAGGCCCTGACCCCACTGTCTCT 1380
DB 11320 CATTCCTTCCAGAGATGTTGAAATGTTCACTCTCCAGGCCCTGACCCCACTGTCTCT 11379
QY 1381 GGAATCAGGAGTGTCTTCCCAACATTTGAGCTGACCGTGTCTCTTAATGAACTCTGG 1440
DB 11380 GGAATCAGGAGTGTCTTCCCAACATTTGAGCTGACCGTGTCTCTTAATGAACTCTGG 11439
QY 1441 AACATTTCCAAACTGTTCAGAGGCGGGGTTGCTCTCAATCTCCCTGGGCACTTTCA 1500
DB 11440 AACATTTCCAAACTGTTCAGAGGCGGGGTTGCTCTCAATCTCCCTGGGCACTTTCA 11499
QY 1501 TCTCTAGGCTCAGGCGCCATCCCTTCTGTGAGCTGTGACCCCAATTTAGTCCAGAAAT 1560
DB 11500 TCTCTAGGCTCAGGCGCCATCCCTTCTGTGAGCTGTGACCCCAATTTAGTCCAGAAAT 11559
QY 1561 AAATGAGAG 1571
DB 11560 AAATGAGAG 11570

RESULT 5
ACH74409
ID ACH74409 standard; DNA; 567 BP.
XX
AC ACH74409;
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #7604.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
OS Homo sapiens.
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
PI Penn SG, Rank DR, Hanzel DK;
PI
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 7604; 80bp; English.
XX

CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of seiling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscrition, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. Note: The sequence data for this
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704

XX Sequence 567 BP; 105 A; 197 C; 125 G; 140 T; 0 U; 0 Other;

Query Match 34.4%; Score 541; DB 12; Length 567;

Best Local Similarity 100.0%; Pred. No. 7e-128;

Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1031 GTGCTCTTGGCCCACTTTGGCAATCTCTGCAATGAGCCCTTCTCTCTGCG 1090
Db 1 GTCTCTTCTGCCCCACTTTGGCAATCTCTGCAATGAGCCCTTCTCTCTGCG 60
Qy 1091 AGGGTATTCCTGGGGGCTGTGTGTGCAATGAGCTCTCTGCAAGGACTGTCTCTGGG 1150
Db 61 AGGGTATTCCTGGGGGCTGTGTGTGCAATGAGCTCTCTGCAAGGACTGTCTCTGGG 120
Qy 1151 GAGATTACCTTGTGCCCCGCGCAACGAGCGGGTGTCTACGCAACCTTGTGCAAGTTCA 1210
Db 121 GAGATTACCTTGTGCCCCGCGCAACGAGCGGGTGTCTACGCAACCTTGTGCAAGTTCA 180
Qy 1211 CCAAGTGGATTCAGGAAGAACATTCAGGCGCAACTCTGATGATCTCCAGAGCTGACAGC 1270
Db 181 CCAAGTGGATTCAGGAAGAACATTCAGGCGCAACTCTGATGATCTCCAGAGCTGACAGC 240
Qy 1271 CGGATCCCACTGTCTGCAAGGAGCAAGCCCTGACATCTCTTTCAGACCTTATCTCTTC 1330
Db 241 CGGATCCCACTGTCTGCAAGGAGCAAGCCCTGACATCTCTTTCAGACCTTATCTCTTC 300
Qy 1331 CAGAGATGTTGAGAAATTTATCTCTCAGGCCCTGACCCCATGCTCTCTGAGCTCAGGG 1390
Db 301 CAGAGATGTTGAGAAATTTATCTCTCAGGCCCTGACCCCATGCTCTCTGAGCTCAGGG 360
Qy 1391 TCTGCTTCCCCCAATTTGGGCTGACCGATGCTCTCTAGTTGGAACCTGGGAAATTTCC 1450
Db 361 TCTGCTTCCCCCAATTTGGGCTGACCGATGCTCTCTAGTTGGAACCTGGGAAATTTCC 420
Qy 1451 AAAAATGTTCAGAGGCGGGGGTGTGTGTCTCAATCTCTCTGGGGCACTTTCATCTCAAGCT 1510
Db 421 AAAAATGTTCAGAGGCGGGGGTGTGTGTCTCAATCTCTCTGGGGCACTTTCATCTCAAGCT 480

Qy 1511 CAGGGCCCATCTCTCTGCGAGCTCTGACCCCAATTTAGTCCCAAGAAATTAAGTGAAG 1570
Db 481 CAGGGCCCATCTCTCTGCGAGCTCTGACCCCAATTTAGTCCCAAGAAATTAAGTGAAG 540
Qy 1571 G 1571
Db 541 G 541
RESULT 6
ABL63579/c
ID ABL63579 standard; DNA; 586 BP.
XX
AC ABL63579;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1916.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN NC0200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001NC-05010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 18-SEP-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0231313P.
PR 20-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 28-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236843P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.

PR 01-NOV-2000; 2000US-024867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX
PA (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a chemical
PT agent to be tested for anti-neoplastic activity, and determining a change
PT in expression of a gene of a signature gene set.
XX
XX
PS Claim 1; SEQ ID NO 1916; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
CC neoplastic agent. The method involves exposing cells to a chemical agent
CC to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I1) of a signature gene set, where (I1)
CC comprises a sequence (S1) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S1), where a change in
CC expression is indicative of anti-neoplastic activity. (I1) has cytoskeletal
CC activity and can be used in gene therapy. M1 can be used for screening an
CC anti-neoplastic agent, and can be used for producing a product which is
CC the data collected with respect to the anti-neoplastic agent as a result
CC of M1, and the data is sufficient to convey the chemical structure and/or
CC properties of the agent. M1 can be used in the treatment of cancer such
CC as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
CC prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
CC cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
CC cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
CC tumour
XX
SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 0 U; 2 Other;
Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 8.3e-113;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGGATTCCTGGGGGCTGTGTCGTAAGGCTCCCTGAGGAGCTCGTCTCGG 1149
DB 501 CAGGGGATTCCTGGGGGCTGTGTCGTAAGGCTCCCTGAGGAGCTCGTCTCGG 442
QY 1150 GGAGATTACCTCTGTCCCGGCCAACAAGACGGGGTGTCTACACGAACCTCTGAGTTTC 1209
DB 441 GGAGATTACCTCTGTCCCGGCCAACAAGACGGGGTGTCTACACGAACCTCTGAGTTTC 382
QY 1210 ACCAAGTGAATCCAGGAACCATTCAGGCCCACTCCTGAGTCAATCCAGAGCTCAGCACA 1269
DB 381 ACCAAGTGAATCCAGGAACCATTCAGGCCCACTCCTGAGTCAATCCAGAGCTCAGCACA 322
QY 1270 CCGGCAATCCCACTCTGTGAGGAGACGCTGACACTCTTTCAAGCCTCAATTCCTTC 1329
DB 321 CCGGCAATCCCACTCTGTGAGGAGACGCTGACACTCTTTCAAGCCTCAATTCCTTC 262
QY 1330 CCAGAGATTTGAGATGTTTCACTCTCCAGCCCTGAGCCCACTGTCCTGAGACTCAGG 1389
DB 261 CCAGAGATTTGAGATGTTTCACTCTCCAGCCCTGAGCCCACTGTCCTGAGACTCAGG 202
QY 1390 GTCTGCTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTTGGGAACAATTC 1449
DB 201 GTCTGCTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTTGGGAACAATTC 142
QY 1450 CAAAATGTCGAGGGGGGGTGTGCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC 1509
DB 141 CAAAATGTCGAGGGGGGGTGTGCTCAATCTCCCTGGGCACTTTTCACTCTCAAGC 82
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
DB 81 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 22
QY 1570 AG 1571

DB 21 AG 20
RESULT 7
ABL63996/C
ID ABL63996 standard; DNA; 586 BP.
XX
XX ABL63996;
AC
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:2333.
XX
XX Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
XX cytostatic; gene therapy; antineoplastic; Wilms tumour; adenocarcinoma;
XX gene; db.
OS Homo sapiens.
XX
XX WO200194629-A2.
PN
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0231133P.
PR 18-SEP-2000; 2000US-023617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 02-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.
PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX
XX (AVAL-) AVALON PHARM.
PA
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;

QY 1450 CAAAGTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
 DB 140 CAAAATGTCTCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 81
 QY 1510 TCAGGGCCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCGGAATAAATGAGA 1569
 DB 80 TCAGGGCCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCGGAATAAATGAGA 21
 QY 1570 AG 1571
 DB 20 AG 19
 RESULT 9
 ID ABK29013/c
 XX ABK29013 standard; cDNA; 735 BP.
 AC ABK29013;
 DT 23-APR-2002 (first entry)
 XX
 DE Human breast tumour polypeptide cDNA clone #42.
 XX
 KW Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
 XX immunostimulant.
 OS Homo sapiens.
 XX
 PN MO200198339-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 12-JUN-2001; 2001WO-US019032.
 XX
 PR 22-JUN-2000; 2000US-00602877.
 PR 12-OCT-2000; 2000US-00687507.
 PR 06-FEB-2001; 2001US-00778381.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
 XX
 DR WPI; 2002-147792/19.
 XX
 PT Polynucleotides encoding breast polypeptides, useful for treating
 PT breast cancer or stimulating an immune response.
 XX
 PS Claim 1; Page 140; 150pp; English.
 XX
 CC The invention relates to polynucleotides encoding breast tumour
 CC polypeptides. The sequences are useful for treating cancer, preferably
 CC breast cancer, in a patient or for stimulating an immune response. The
 CC polynucleotides and polypeptides are also useful in the diagnosis and
 CC monitoring of breast cancer. A method for detecting the presence of a
 CC cancer in a patient, comprises obtaining a biological sample from the
 CC patient, contacting the biological sample with a binding agent that binds
 CC to a breast tumour polypeptide, detecting in the sample an amount of
 CC polypeptide that binds to the binding agent, and comparing the amount of
 CC polypeptide to a predetermined cut-off value, therefore determining the
 CC presence of a cancer in the patient. Sequences ABK28920-ABK29025
 CC represent cDNA clones encoding human breast tumour polypeptides of the
 CC invention
 CC
 SQ Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 U; 0 Other;
 Query Match 30.7%; Score 482; DB 6; Length 735;
 Best Local Similarity 100.0%; Pred. No. 9e-113;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1090 CAGGGGATTTCTGGGGGCTGTGCTGCAATGCTCCCTGAGAGACTCGTCTCGG 1149
 DB 500 CAGGGGATTTCTGGGGGCTGTGCTGCAATGCTCCCTGAGAGACTCGTCTCGG 441

QY 1150 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
 DB 440 GGAGATTACCTTGTGTCGCCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAGTTC 381
 QY 1210 ACCAAGTGTATCAGAGAAACCTTCCAGGCTCAACTCTGATGATTCAGAGACTCAGCA 1269
 DB 380 ACCAAGTGTATCAGAGAAACCTTCCAGGCTCAACTCTGATGATTCAGAGACTCAGCA 321
 QY 1270 CCGGATCCCACTGTCTGACAGAGACAGCCCTGACACTCTTTTCAGACCTCATTCCTTC 1329
 DB 320 CCGGATCCCACTGTCTGACAGAGACAGCCCTGACACTCTTTTCAGACCTCATTCCTTC 261
 QY 1330 CCAGAGATGTTGAGATGTTCTATCTCTCAGCCCTGACCCCATGTTCTTCTGACTCAGG 1389
 DB 260 CCAGAGATGTTGAGATGTTCTATCTCTCAGCCCTGACCCCATGTTCTTCTGACTCAGG 201
 QY 1390 GTCTGCTTCCCAATTTGGGTGACCGTGTCTCTAGTTGAACCTTGGGAACTTTC 1449
 DB 200 GTCTGCTTCCCAATTTGGGTGACCGTGTCTCTAGTTGAACCTTGGGAACTTTC 141
 QY 1450 CAAAGTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1509
 DB 140 CAAAATGTCTCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 81
 QY 1510 TCAGGGCCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCGGAATAAATGAGA 1569
 DB 80 TCAGGGCCCATCTCTCTCTGAGCTCTGACCCCAATTTAGTCCGGAATAAATGAGA 21
 QY 1570 AG 1571
 DB 20 AG 19
 RESULT 10
 ID AAX84240/c
 XX AAX84240 standard; cDNA; 738 BP.
 AC AAX84240;
 XX
 DT 08-SEP-1999 (first entry)
 XX
 DE DNA encoding human breast tumour protein immunogenic fragment.
 XX
 KW Breast tumour protein; immunogenic fragment; vaccine; detection;
 KW breast cancer development; therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9333869-A2.
 XX
 PD 08-JUL-1999.
 XX
 PF 22-DEC-1998; 98WO-US027416.
 XX
 PR 24-DEC-1997; 97US-00998253.
 PR 24-DEC-1997; 97US-00998255.
 PR 17-JUL-1998; 98US-00118554.
 PR 17-JUL-1998; 98US-00118627.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Reed SG, Xu J;
 XX
 DR WPI; 1999-405486/34.
 XX
 PT New breast tumor protein genes used, in vaccines for immunotherapy, or
 PT for diagnosis of breast cancer.
 XX
 PS Claim 3; Page 70; 70pp; English.
 XX
 CC This sequence encodes a human breast tumour protein immunogenic fragment
 CC of the invention. The polypeptides or nucleic acids encoding them are
 CC useful in vaccines and pharmaceutical compositions for manufacture of

medicaments for inhibiting the development of breast cancer in a patient.
CC They can also be used to treat breast cancer. Antibodies against these
CC polypeptides can be used to detect and monitor progression of breast
CC cancer in patients. Primers and probes derived from the polynucleotides
CC encoding the breast proteins are useful for detection of breast cancer.
CC Peripheral blood cells from a patient incubated in the presence of at
CC least one polypeptide, such that T cells proliferate, are useful in
CC manufacture of a medicament for treating breast cancer in a patient.
CC Antigen presenting cells incubated in the presence of at least one
CC polypeptide are also useful for treating breast cancer
XX
SQ Sequence 738 BP; 161 A; 179 C; 226 G; 172 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 2; Length 738;
Best Local Similarity 100.0%; Pred. No. 9.1e-113; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTATTCTGGGGGCGCTGTGTCGATGGCTCCCTGCAGGACTGTCTCTGG 1149
DB 503 CAGGGTATTCTGGGGGCGCTGTGTCGATGGCTCCCTGCAGGACTGTCTCTGG 444
QY 1150 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACAGAACTTCTGCAAGTTC 1209
DB 443 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACAGAACTTCTGCAAGTTC 384
QY 1210 ACCAAGTGTGATCCAGGAAACCATTCAGGCCCACTCTGAGTCACTCCAGACTCAGACACA 1269
DB 383 ACCAAGTGTGATCCAGGAAACCATTCAGGCCCACTCTGAGTCACTCCAGACTCAGACACA 324
QY 1270 CCGGCATCCCACTGCTGTCAGGAGAGCCCTGACATCCTTTCAGACCCCTCATTCCTTC 1329
DB 323 CCGGCATCCCACTGCTGTCAGGAGAGCCCTGACATCCTTTCAGACCCCTCATTCCTTC 264
QY 1330 CCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1389
DB 263 CCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 204
QY 1390 GTGTGCTTCCCACTGTCAGTGGCTGACCGTGTCTCTGATGTAACCTTGGGAAACAATTC 1449
DB 203 GTGTGCTTCCCACTGTCAGTGGCTGACCGTGTCTCTGATGTAACCTTGGGAAACAATTC 144
QY 1450 CAAAAGTGTCCAGGGGCGGGGTGCGTCTCAATCTCCCTGGGAGCACTTTCATCTCAAGC 1509
DB 143 CAAAAGTGTCCAGGGGCGGGGTGCGTCTCAATCTCCCTGGGAGCACTTTCATCTCAAGC 84
QY 1510 TCAAGGCGCCATCCCTTCTCTGACAGCTGACCCAAATTAAGTCCAGAAATMAACTGAGA 1569
DB 83 TCAAGGCGCCATCCCTTCTCTGACAGCTGACCCAAATTAAGTCCAGAAATMAACTGAGA 24
QY 1570 AG 1571
DB 23 AG 22
RESULT 11
ADB80489 standard; DNA; 1260 BP.
XX
XX ADB80489;
DT 04-DEC-2003 (first entry)
XX
XX Ovarian cancer-associated transcript #27.
XX
XX cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection; ds; gene.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 43..924
FT CDS
FT /*tag= a

XX
PN WO2002102235-A2.
XX
XX 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
PR 27-AUG-2001; 2001US-0315287P.
PR 05-SEP-2001; 2001US-0317544P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Mack DH, Gish KC;
PI
DR WPI; 2003-167431/16.
XX
DR P-PSDB; ADB80490.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS Claim 10; Page 292; 332pp; English.
XX
CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancers, monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the nucleic acids
CC used for the detection method of the invention.
XX
SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;
Query Match 30.7%; Score 482; DB 10; Length 1260;
Best Local Similarity 100.0%; Pred. No. 1.1e-112; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTATTCTGGGGGCGCTGTGTCGATGGCTCCCTGCAGGACTGTCTCTGG 1149
DB 766 CAGGGTATTCTGGGGGCGCTGTGTCGATGGCTCCCTGCAGGACTGTCTCTGG 825
QY 1150 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACAGAACTTCTGCAAGTTC 1209
DB 826 GGAAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTACAGAACTTCTGCAAGTTC 885
QY 1210 ACCAAGTGTGATCCAGGAAACCATTCAGGCCCACTCTGAGTCACTCCAGACTCAGACACA 1269
DB 886 ACCAAGTGTGATCCAGGAAACCATTCAGGCCCACTCTGAGTCACTCCAGACTCAGACACA 945
QY 1270 CCGGCATCCCACTGCTGTCAGGAGAGCCCTGACATCCTTTCAGACCCCTCATTCCTTC 1329
DB 946 CCGGCATCCCACTGCTGTCAGGAGAGCCCTGACATCCTTTCAGACCCCTCATTCCTTC 1005
QY 1330 CCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1389
DB 1006 CCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGACTGAG 1065
QY 1390 GTGTGCTTCCCACTGTCAGTGGCTGACCGTGTCTCTGATGTAACCTTGGGAAACAATTC 1449
DB 1066 GTGTGCTTCCCACTGTCAGTGGCTGACCGTGTCTCTGATGTAACCTTGGGAAACAATTC 1125
QY 1450 CAAAAGTGTCCAGGGGCGGGGTGCGTCTCAATCTCCCTGGGAGCACTTTCATCTCAAGC 1509

Db	1126	CANAACGTGTCAGAGGCGGGGGTGGTGTCAATCTCCCTGGGAGCACTTATCTCAAGC	1185
Qy	1510	TCAGGGCCCATCCCTTCTCTGACAGCTTGACCCAAATTAGTCCAGAAATPAACTGAGA	1569
Db	1186	TCAGGGCCCATCCCTTCTCTGACAGCTTGACCCAAATTAGTCCAGAAATPAACTGAGA	1245
Qy	1570	AG 1571	
Db	1246	AG 1247	
RESULT 12			
ID	ADN39197		
AC	ADN39197	standard; cDNA, 1260 BP.	
XX	ADN39197;		
DT	17-JUN-2004	(first entry)	
DE	Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.		
XX			
KW	Human; differential expression; cancer; angiogenic disorder;		
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;		
KW	inflammatory disease; autoimmune disease;		
KW	retinal neovascularisation syndrome; scarring; uterine fibroid;		
KW	detection; diagnosis; prognosis; drug screening; drug targeting;		
KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;		
KW	vulnerable; gene therapy; vaccine; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003042661-A2.		
PD	22-MAY-2003.		
XX			
PF	13-NOV-2002; 2002WO-US036810.		
XX			
PR	13-NOV-2001; 2001US-0350666P.		
PR	21-NOV-2001; 2001US-0332464P.		
PR	29-NOV-2001; 2001US-0334393P.		
PR	03-DEC-2001; 2001US-0335394P.		
PR	14-DEC-2001; 2001US-0340376P.		
PR	08-JAN-2002; 2002US-0347231P.		
PR	10-JAN-2002; 2002US-0347349P.		
PR	08-FEB-2002; 2002US-035250P.		
PR	13-FEB-2002; 2002US-0356714P.		
PR	20-FEB-2002; 2002US-0359077P.		
PR	29-MAR-2002; 2002US-0368809P.		
PR	04-APR-2002; 2002US-0370110P.		
PR	12-APR-2002; 2002US-0372246P.		
PR	05-JUN-2002; 2002US-0386614P.		
PR	16-JUL-2002; 2002US-0396839P.		
PR	22-JUL-2002; 2002US-0397775P.		
PR	22-JUL-2002; 2002US-0397845P.		
PR	09-SEP-2002; 2002US-0409450P.		
XX			
PA	(EOSB-) EOS BIOTECHNOLOGY INC.		
XX			
PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glyne R, Hevezi PA;		
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;		
XX			
DR	WPI; 2003-468649/44.		
XX	P-PSDB; ADN39198.		
PT			
PT	Determining the presence or absence of a pathological cell in a patient,		
XX	useful for diagnosing, prognosing or treating cancer, comprises detecting		
XX	a nucleic acid in a biological sample.		
XX			
XX	Claim 8; SEQ ID NO 515; 1385bp; English.		
CC			
CC	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)		
CC	whose expression is upregulated or downregulated in specific cancers or		
CC	other diseases such as angiogenic or fibrotic disorders, and to methods		

CC		of determining the presence or absence of a pathological cell in a
CC		patient by detecting a nucleic acid at least 80% identical to those of
CC		the invention or by detecting a polypeptide of the invention. The
CC		invention also relates to expression vectors and host cells comprising a
CC		nucleic acid of the invention; antibodies which specifically bind a
CC		polypeptide of the invention; use of such antibodies for drug targeting;
CC		and methods of screening for modulators of activity or expression of the
CC		polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC		antibodies and methods are useful for diagnosing, prognosing and treating
CC		cancer and other conditions such as psoriasis, leukaemia, heart disease,
CC		atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC		neovascularisation syndromes, scarring and uterine fibroids. The present
CC		also be useful in wound healing and in contraception. The present
CC		sequence represents a nucleic acid sequence of the invention.
XX		
SQ	Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;	
	Query Match	30.7%; Score 482; DB 11; Length 1260;
	Best Local Similarity	100.0%; Pred. No. 1.1e-112;
	Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1090 CAGGGTGAATTCGTGGGAGGCGCTGTGGTGTCGACAAATGCTCCCTGCAGGACTGTGTCCTGG	1149
Db	766 CAGGGTGAATTCGTGGGAGGCGCTGTGGTGTCGACAAATGCTCCCTGCAGGACTGTGTCCTGG	825
OY	1150 GGAATTACCTTTGTGCCCGGCCCAACAAGACCGGGGTGTCTAACAAGAACCTTGTCAAGTTC	1209
Db	826 GGAATTACCTTTGTGCCCGGCCCAACAAGACCGGGGTGTCTAACAAGAACCTTGTCAAGTTC	885
OY	1210 ACCAAGTGGATCCAGGAAACCATCCAGGCGCAACTCTGAAGTCATCCAGAGCTCAGACA	1269
Db	886 ACCAAGTGGATCCAGGAAACCATCCAGGCGCAACTCTGAAGTCATCCAGAGCTCAGACA	945
OY	1270 CCGGCATCCCCCACCTGTGTGACGAGGACAGCCCTTGACACTCTTTTGAACCCCTCATTTCTTC	1329
Db	946 CCGGCATCCCCCACCTGTGTGACGAGGACAGCCCTTGACACTCTTTTGAACCCCTCATTTCTTC	1005
OY	1330 CCAGAAGATGTTGAGAAATGTTTCATCTTCACGCCCCCTGAACCCCATGTCTCTGAACTCAGG	1389
Db	1006 CCAGAAGATGTTGAGAAATGTTTCATCTTCACGCCCCCTGAACCCCATGTCTCTGAACTCAGG	1065
OY	1390 GTCTGCTTCCCCTCCCATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAAACAATTTC	1449
Db	1066 GTCTGCTTCCCCTCCCATTTGGGCTGACCGGTGTCTCTAGTTGAACCCCTGGGAAACAATTTC	1125
OY	1450 CAATAAGTGTCCAGGCGCGGGGTTGCGTCTCAATCTCCCTGGGAGCACTTTCATCTCAAGC	1509
Db	1126 CAATAAGTGTCCAGGCGCGGGGTTGCGTCTCAATCTCCCTGGGAGCACTTTCATCTCAAGC	1185
OY	1510 TCAGGGGCCCATCCCTTCTCTGACGCTTGACCCAAATTTAGTCCCAAGAAATTAACCTGAGA	1569
Db	1186 TCAGGGGCCCATCCCTTCTCTGACGCTTGACCCAAATTTAGTCCCAAGAAATTAACCTGAGA	1245
OY	1570 AG 1571	
Db	1246 AG 1247	
	RESULT 13	
ID	ADR25550	
1D	ADR25550 standard; DNA; 1260 BP.	
XX	ADR25550;	
DT	21-OCT-2004 (first entry)	
XX	Breast cancer prognosis marker #1411.	
DE	de; breast cancer; prognosis; gene expression; diagnosis.	
XX	Homo sapiens	
PN	WO2004065545-A2.	

QY 1330 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCAATGTCCTCGACTCAGG 1389
|||
DB 1051 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCAATGTCCTCGACTCAGG 1110
|||
QY 1390 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCGGAACAATTC 1449
|||
DB 1111 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCGGAACAATTC 1170
|||
QY 1450 CAAACATGTCAGAGGGGGGCTTGCCTTCATCTCCCTGGGCACTTTCATCTCAAG 1509
|||
DB 1171 CAAACATGTCAGAGGGGGGCTTGCCTTCATCTCCCTGGGCACTTTCATCTCAAG 1230
|||
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1569
|||
DB 1231 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1290
|||
QY 1570 AG 1571
|||
DB 1291 AG 1292

RESULT 15

AAK57989 standard; DNA; 1302 BP.

AAK57989;

19-JUL-1999 (first entry)

Human BS247 specific polynucleotide #7.

BS247; detection; diagnosis; breast cancer; atypical hyperplasia;
fibroadenoma; cystic breast disease; gene therapy; ss.

Homo sapiens.

MO9222027-A1.

06-MAY-1999.

28-OCT-1998; 98MO-US022906.

28-OCT-1997; 97US-0063431P.

28-OCT-1997; 97US-00968838.

(ABBO) ABBOT LAB.

Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
Granados EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC;
Stroupe SD, Yu H;

WPI; 1999-312977/26.

Breast tissue derived cDNA contig and consensus polypeptide sequence.

Claim 1; Page 105; 112pp; English.

This sequence is a BS247 specific polynucleotide. The invention relates to a method of detecting the presence of a target BS247 polynucleotide, especially mRNA, in a test sample. BS247 polynucleotides are derived from breast tissue. The polynucleotides, polypeptides or antibodies are useful for providing information leading to the detection, diagnosis, staging, monitoring, prognosis, in vivo imaging, prevention or treatment, determining predisposition to, diseases and conditions of the breast, such as breast cancer, atypical hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or gene therapy for breast cancer, can be based on these identified gene sequences and the efficacy of any particular therapy can be monitored. The BS247-derived reagents are advantageous for detection of breast cancer due to their specificity. The reagents also provide an alternative, non-surgical diagnostic method capable of detecting early stage breast disease, such as cancer

SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;

Query Match 30.7%; Score 482; DB 2; Length 1302;

Best Local Similarity 100.0%; Pred. No. 1.1e-112;

Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGATGATTTGGGGGGGCTGTGTGTCGCAATGCTCCCTGACAGGACTCGTGTCTGG 1149
|||
DB 811 CAGGATGATTTGGGGGGGCTGTGTGTCGCAATGCTCCCTGACAGGACTCGTGTCTGG 870
|||
QY 1150 GGAGATTAACCTTTGTGCGCGCCCAACAGACCGGGTGTACAGAACTCTGCAATTC 1209
|||
DB 871 GGAGATTAACCTTTGTGCGCGCCCAACAGACCGGGTGTACAGAACTCTGCAATTC 930
|||
QY 1210 ACCAATGATTCAGAGAAACCATCGAGGCAATCTGAGTATCCAGAGACTCAGACA 1269
|||
DB 931 ACCAATGATTCAGAGAAACCATCGAGGCAATCTGAGTATCCAGAGACTCAGACA 990
|||
QY 1270 CCGGATCCCACTGCTGACAGGACAGCCCTGACATCTCTTTCAGACCTTCATTCCTTC 1329
|||
DB 991 CCGGATCCCACTGCTGACAGGACAGCCCTGACATCTCTTTCAGACCTTCATTCCTTC 1050
|||
QY 1330 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCAATGTCCTCGACTCAGG 1389
|||
DB 1051 CCAGAGATGTTGAGAAATGTCATCTCTCCAGCCCTGAGCCCAATGTCCTCGACTCAGG 1110
|||
QY 1390 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCGGAACAATTC 1449
|||
DB 1111 GTCTGCTTCCCAACATTTGGGCTGACCGTGTCTCTAGTTGAACCTCGGAACAATTC 1170
|||
QY 1450 CAAACATGTCAGAGGGGGGCTTGCCTTCATCTCCCTGGGCACTTTCATCTCAAG 1509
|||
DB 1171 CAAACATGTCAGAGGGGGGCTTGCCTTCATCTCCCTGGGCACTTTCATCTCAAG 1230
|||
QY 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1569
|||
DB 1231 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAATGAGA 1290
|||
QY 1570 AG 1571
|||
DB 1291 AG 1292

Search completed: February 25, 2005, 20:15:07

Job time : 802.397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 18:05:22 : Search time 6692.88 Seconds
(without alignments)
11373.754 Million cell updates/sec

Title: US-09-936-271B-13_COPY_10000_11570

Perfect score: 1571
Sequence: 1 agggagagctctgggaactga.....ccaggaataactgagaag 1571

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Genbank: *
1: gb_ba: *
2: gb_hg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_dl: *
9: gb_pr: *
10: gb_ro: *
11: gb_stb: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	1571	100.0	11570	6	CQ788219	CQ788219 Sequence
2	1571	100.0	11570	6	CQ874885	CQ874885 Sequence
3	1571	100.0	11570	6	CQ874961	CQ874961 Sequence
4	1571	100.0	11570	9	AF135028	AF135028 Homo sapi
5	1571	100.0	107487	9	AF011483	AF011483 Homo sapi
6	1571	100.0	230000	9	AF243527	AF243527 Homo sapi
7	1555.2	99.0	217346	2	AC027602	AC027602 Homo sapi
8	1482	94.3	200792	2	AC130782	AC130782 Pan trogl
9	482	30.7	586	6	AX331407	AX331407 Sequence
10	482	30.7	586	6	AX331407	AX331407 Sequence
11	482	30.7	586	6	AX331824	AX331824 Sequence
12	482	30.7	735	6	AX429955	AX429955 Sequence
13	482	30.7	738	6	BD139877	BD139877 Compound
14	482	30.7	1143	6	CQ720600	CQ720600 Sequence
15	482	30.7	1301	9	AY461805	AY461805 Homo sapi
16	482	30.7	1370	9	AY279380	AY279380 Homo sapi
17	482	30.7	1381	6	BD107879	BD107879 36 human
18	482	30.7	1387	6	CQ874884	CQ874884 Sequence
19	482	30.7	1387	6	CQ874960	CQ874960 Sequence

20	482	30.7	1387	9	AF168768	AF168768 Homo sapi
21	482	30.7	1438	6	AY279381	AY279381 Homo sapi
22	482	30.7	1439	6	BD107880	BD107880 36 human
23	482	30.7	1439	6	AR352504	AR352504 Sequence
24	482	30.7	1499	6	BD005362	BD005362 Protease
25	482	30.7	1516	6	BD107865	BD107865 36 human
26	482	30.7	1527	6	BC008036	BC008036 Homo sapi
27	482	30.7	1570	6	AR252580	AR252580 Sequence
28	482	30.7	1570	6	AR528679	AR528679 Sequence
29	482	30.7	1570	6	AX080829	AX080829 Sequence
30	482	30.7	1570	6	AX403421	AX403421 Sequence
31	482	30.7	1570	6	AX464322	AX464322 Sequence
32	482	30.7	1570	6	AY359010	AY359010 Homo sapi
33	476	30.3	1476	6	AR078184	AR078184 Sequence
34	476	30.3	1476	6	AR137506	AR137506 Sequence
35	476	30.3	1476	6	AR242358	AR242358 Sequence
36	476	30.3	1476	6	BD082659	BD082659 Keratinoc
37	460	29.3	1504	6	AR263823	AR263823 Sequence
38	449	28.6	673	6	AX067353	AX067353 Sequence
39	315.4	20.1	498	11	BV209338	BV209338 KLRK5_3564
40	236.6	15.1	321519	2	AL714004	AL714004 Homo sapi
41	236	15.0	175559	9	AC112458	AC112458 Homo sapi
42	234.4	14.9	174380	2	AC016075	AC016075 Homo sapi
43	233.2	14.8	151880	9	AP005716	AP005716 Homo sapi
44	231.2	14.7	79516	9	AC004834	AC004834 Homo sapi
45	231.2	14.7	156949	9	AL161652	AL161652 Human DNA

ALIGNMENTS

RESULT 1	CQ788219	CQ788219	11570 bp	DNA	linear	PAT 24-MAR-2004
LOCUS	Sequence 2 from Patent WO2004021008.					
DEFINITION	CQ788219					
ACCESSION	CQ788219					
VERSION	CQ788219.1	GI:45723068				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1	Diamonds, E. P.				
AUTHORS		Methods for detecting breast and ovarian cancer				
TITLE		Patent: WO 2004021008-A 2 11-MAR-2004;				
JOURNAL		MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)				
FEATURES		Location/Qualifiers				
source		1..11570				
ORIGIN		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				
Query Match	100.0%	Score 1571;	DB 6;	Length 11570;		
Best Local Similarity	100.0%	Pred. No. 0;				
Matches 1571;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
QY	1	AGGAGGATGGGGAATTGAGACAGAAACAAATTAAGTCCAGCGATGATTCTA	60			
DB	10000	AGGAGGATGGGGAATTGAGACAGAAACAAATTAAGTCCAGCGATGATTCTA	10059			
QY	61	TTGGAGTGTATTCCTGCTTGAAGACATGCGAATACCAAGAGACATTTTGGTTGTCA	120			
DB	10060	TTGGAGTGTATTCCTGCTTGAAGACATGCGAATACCAAGAGACATTTTGGTTGTCA	10119			
QY	121	CAACTATATGAGGGGCACTTACGCACTAATGATGATGCAAGTGGCTTCAAC	180			
DB	10120	CAACTATATGAGGGGCACTTACGCACTAATGATGATGCAAGTGGCTTCAAC	10179			
QY	181	ATGCTATGATGACACAGCGAGGCTTCACAAACCAATATTCAGATGCCAC	240			
DB	10180	ATGCTATGATGACACAGCGAGGCTTCACAAACCAATATTCAGATGCCAC	10239			

QY 241 AGTCCAGATCGAGGAACTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGAGGAG 300
Db 10240 AGTGCCAGATGAGAGAACCTTCATCCAGGGCTGAGAACCGTATTTTTCAGAGAGGAG 10299
QY 301 GTATTAAGATGGTGTGTGTGAGAAATGGGAGAGAGGTGTGTGTCCATGAGAAATTA 360
Db 10300 GTATTAAGATGGTGTGTGTGAGAAATGGGAGAGAGGTGTGTGTCCATGAGAAATTA 10359
QY 361 GGCCTGACAGAGCTGAGAGGGGAGAGTGAAGAGAGAGAGGGGAGAGAGTACAGATTA 420
Db 10360 GGCCTGACAGAGCTGAGAGGGGAGAGTGAAGAGAGAGAGGGGAGAGAGTACAGATTA 10419
QY 421 GGGAGA CAGGCTGAGAACAGAAATAGAGAGAGATTTCCAGATGTGAGAGAGAGGCTCA 480
Db 10420 GGGAGA CAGGCTGAGAACAGAAATAGAGAGAGATTTCCAGATGTGAGAGAGAGGCTCA 10479
QY 481 CAGACCCCCCGGAATGATGTGTGACAA CAGGAATCTGAGAGAGAGAGATGAGATGAG 540
Db 10480 CAGACCCCCCGGAATGATGTGTGACAA CAGGAATCTGAGAGAGAGAGATGAGATGAG 10539
QY 541 AGTGAACAATGGGGTCTAAAGGTGAACCTTGGAGGCTGAGGAGATGTGCTCAGCCTGTA 600
Db 10540 AGTGAACAATGGGGTCTAAAGGTGAACCTTGGAGGCTGAGGAGATGTGCTCAGCCTGTA 10599
QY 601 ATCCCAACACTTGGAGGCTGAGAGGTGGGGAATCACTTGGAGGCTGAGAGATTCAGAGACAG 660
Db 10600 ATCCCAACACTTGGAGGCTGAGAGGTGGGGAATCACTTGGAGGCTGAGAGATTCAGAGACAG 10659
QY 661 CCTGGCAACATGTGTGAACCCCGTCTCTCAAAAAAAATTAAGCCGGGTGT 720
Db 10660 CCTGGCAACATGTGTGAACCCCGTCTCTCAAAAAAAATTAAGCCGGGTGT 10719
QY 721 GGTATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGAGAGATGTGCTGAGAC 780
Db 10720 GGTATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGAGAGATGTGCTGAGAC 10779
QY 781 CCGGAGATGAGAGGCTGAGAGTGAAGGTCAAGGCACTGCGCTCAACCTGGGCAACA 840
Db 10780 CCGGAGATGAGAGGCTGAGAGTGAAGGTCAAGGCACTGCGCTCAACCTGGGCAACA 10839
QY 841 GAGTAAGATCTCCATCTCAAAAAAAAGCTGAGATTTGAGAGTAATTAATTAACATT 900
Db 10840 GAGTAAGATCTCCATCTCAAAAAAAAGCTGAGATTTGAGAGTAATTAATTAACATT 10899
QY 901 CTCCTCTCTCTCTTGTGCTGTGTCTCCATCTGCTTTTTCGACATTTCTTCACT 960
Db 10900 CTCCTCTCTCTCTTGTGCTGTGTCTCCATCTGCTTTTTCGACATTTCTTCACT 10959
QY 961 CTGTACTTTCATCTGTGTGTCTGTCTCCATCTGCTTTTTCGACATTTCTTCACT 1020
Db 10960 CTGTACTTTCATCTGTGTGTCTGTCTCCATCTGCTTTTTCGACATTTCTTCACT 11019
QY 1021 GGTCTCTCATGTCTCTTCTGCCCACCTTGGCAACATCTGCTGCTCTCAATGCCCTTT 1080
Db 11020 GGTCTCTCATGTCTCTTCTGCCCACCTTGGCAACATCTGCTGCTCTCAATGCCCTTT 11079
QY 1081 TCTCTCTGAGAGGATTCGTGGGGGCGTGTGTGCAATGGGCTCCGCAAGGAGAC 1140
Db 11080 TCTCTCTGAGAGGATTCGTGGGGGCGTGTGTGCAATGGGCTCCGCAAGGAGAC 11139
QY 1141 GTGTCTGAGAGATTAACCTTGTGCCCGGCCCAACAGACCGGGGTCTTCAAGAACCTC 1200
Db 11140 GTGTCTGAGAGATTAACCTTGTGCCCGGCCCAACAGACCGGGGTCTTCAAGAACCTC 11199
QY 1201 TGCAGATTCAACAGATGATTCAGAGAACCATTCAGGCAACTCTGAGTCAATCCAGGA 1260
Db 11200 TGCAGATTCAACAGATGATTCAGAGAACCATTCAGGCAACTCTGAGTCAATCCAGGA 11259
QY 1261 CTCAGACACCGGCAATCCCACTGCTGAGAGGAGACCCCTGACATCTCTTTCAGACCCCT 1320
Db 11260 CTCAGACACCGGCAATCCCACTGCTGAGAGGAGACCCCTGACATCTCTTTCAGACCCCT 11319

QY 1321 CATTCCTCCAGAGATGTGAGAAATGTTCATCTCTCAGGCCCTGACCCCATGTCTCT 1380
Db 11320 CATTCCTCCAGAGATGTGAGAAATGTTCATCTCTCAGGCCCTGACCCCATGTCTCT 11379
QY 1381 GGAATCAGGGGTGTCTTCCCTTCCCAATTTGGGCTGACCCGTCTCTTGTGAAACCTTGG 1440
Db 11380 GGAATCAGGGGTGTCTTCCCTTCCCAATTTGGGCTGACCCGTCTCTTGTGAAACCTTGG 11439
QY 1441 AACCAATTTCCAAACGTCTCAGGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCA 1500
Db 11440 AACCAATTTCCAAACGTCTCAGGGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCA 11499
QY 1501 TCCCTAGCTCAGGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTAATGTCAGAAAT 1560
Db 11500 TCCCTAGCTCAGGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTAATGTCAGAAAT 11559
QY 1561 AACTGAGAG 1571
Db 11560 AACTGAGAG 11570

RESULT 2
CQ874885
LOCUS CQ874885 11570 bp DNA linear PAT 27-SEP-2004
DEFINITION Sequence 3 from Patent WO2004077060.
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 Diamandis, E.P. and Petraki, C.
TITLE Assay for detection of renal cell carcinoma
JOURNAL Patent: WO 2004077060-A 3 10-SEP-2004;
Mount Sinai Hospital (CA)
FEATURES
source location/Qualifiers
1..11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1571; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGTAATGGGAATTGAGACAGAAACAAATTAATGTCAGAGGATGATTTCTA 60
Db 10000 AGGAGGTAATGGGAATTGAGACAGAAACAAATTAATGTCAGAGGATGATTTCTA 10059
QY 61 TTGGAGATGATTTCTGCCCCCTAGAGACACTGGCAATACAGAGACATTTTGTGTGCA 120
Db 10060 TTGGAGATGATTTCTGCCCCCTAGAGACACTGGCAATACAGAGACATTTTGTGTGCA 10119
QY 121 CAATATATGAGAGGGGCAATTAATGCACTAATGATAGATGCAAGTGTGCTTCAAC 180
Db 10120 CAATATATGAGAGGGGCAATTAATGCACTAATGATAGATGCAAGTGTGCTTCAAC 10179
QY 181 ATGCTATGATGACACGGGAGGCTTCAACAAACCAATTAATGCTGAGATGATGCCAC 240
Db 10180 ATGCTATGATGACACGGGAGGCTTCAACAAACCAATTAATGCTGAGATGATGCCAC 10239
QY 241 AGTGCCAGATGAGAGAACCTTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG 300
Db 10240 AGTGCCAGATGAGAGAACCTTCATCCAGGGGCTGAGAACCGTATTTTTCAGAGGGAG 10299
QY 301 GTATTAAGATGGTGTGTGTGAGAAATGGGAGAGAGGTGTGTGTCAGTAAAGAAATTA 360
Db 10300 GTATTAAGATGGTGTGTGTGAGAAATGGGAGAGAGGTGTGTGTCAGTAAAGAAATTA 10359
QY 361 GGCCTGACAGGCTGAGAGGGGAGAGTGAAGAGAGAGAGGCGGAGAGATTAACAGATGA 420

Db	10360	GGGCTGCACAGGCTGGAGGGGAGGTGAGAGAGAAAGGAGGGCGAGAGATTACAGATTA	10419
Oy	421	GGGAGACAGGCTGGAAACGAAAGTGAAGACGAAATTCCAGATGTGGAGAGAAAGGGTCA	480
Db	10420	GGGAGACAGGCTGGAAACGAAAGTGAAGAGAAATTCCAGATTGTGGAGAGAAAGGGTCA	10479
Oy	481	CAGACCCCCCGGAAATGATGTGTGGACAACGAAATCTGGAAAGAGAAATGAGTGGAG	540
Db	10480	CAGACCCCCCGGAAATGATGTGTGGACAACGAAATCTGGAAAGAGAAATGAGTGGAG	10539
Oy	541	AGTACAAATGGGGGTCTAAAGTTGAACCTTGGAGGGCAGGACAGTGGGTCAACGCTGTGA	600
Db	10540	AGTACAAATGGGGGTCTAAAGTTGAACCTTGGAGGGCAGGACAGTGGGTCAACGCTGTGA	10599
Oy	601	ATCCCAACACTTTGAGAGGCTGAGGGTGGGCGCAATCACTTGGAGGCGACAGAGTTGACAGACAG	660
Db	10600	ATCCCAACACTTTGAGAGGCTGAGGGTGGGCGCAATCACTTGGAGGCGACAGAGTTGACAGACAG	10659
Oy	661	CTGGGCGCAACATGGTGAACCCCGGTCTTACAAAAAAAATACAAAAAATTAAGCCGGGTGT	720
Db	10660	CTGGGCGCAACATGGTGAACCCCGGTCTTACAAAAAAAATACAAAAAATTAAGCCGGGTGT	10719
Oy	721	GGTATAGGACACTGTATGTACACAGCTACTTGGAGGCTGAGGGAGAGAAATTGCTTGAAC	780
Db	10720	GGTATAGGACACTGTATGTACACAGCTACTTGGAGGCTGAGGGAGAGAAATTGCTTGAAC	10779
Oy	781	CCGGAGATAGGAGGCTGACGTAGAGCTGAGGCCACCTGGCCTCAACCTGGGCAACA	840
Db	10780	CCGGAGATAGGAGGCTGACGTAGAGCTGAGGCCACCTGGCCTCAACCTGGGCAACA	10839
Oy	841	GAGTAAAGACTTCATCTCAAAAAAATAAAGCTGGAATTTGGAGTGAATAATTAATTAACATT	900
Db	10840	GAGTAAAGACTTCATCTCAAAAAAATAAAGCTGGAATTTGGAGTGAATAATTAATTAACATT	10899
Oy	901	CTCCCTCTCTCTCCTTTTGGCTGTGTCTCATCTCTGTCTTTTCTTGCAATTTCTTCATCT	960
Db	10900	CTCCCTCTCTCTCCTTTTGGCTGTGTCTCATCTCTGTCTTTTCTTGCAATTTCTTCATCT	10959
Oy	961	CTTGACTTTTCCATCTCTGTGTGTCTGTCTCCCATCTGGCTTTCATCTAATAGGACATCTGTG	1020
Db	10960	CTTGACTTTTCCATCTCTGTGTGTCTGTCTCCCATCTGGCTTTCATCTAATAGGACATCTGTG	11019
Oy	1021	GGTCTCTCATGTCTCTCTTGCCCACTTTGCAACATCTGCTCTCTCATGTGCCCTTT	1080
Db	11020	GGTCTCTCATGTCTCTCTTGCCCACTTTGCAACATCTGCTCTCTCATGTGCCCTTT	11079
Oy	1081	TCTCTCTGCAAGGGTGAATTCTGGGGGGGCTGTGTGTGTGCAATGGCTTCTGCAAGGACCTC	1140
Db	11080	TCTCTCTGCAAGGGTGAATTCTGGGGGGGCTGTGTGTGTGCAATGGCTTCTGCAAGGACCTC	11139
Oy	1141	GTGTCTGGGGAGATTACCTTTGTGCCGGGCGCAACAGACCGGGTGTCTACACGAACCTTC	1200
Db	11140	GTGTCTGGGGAGATTACCTTTGTGCCGGGCGCAACAGACCGGGTGTCTACACGAACCTTC	11199
Oy	1201	TGCAAGTTTCAACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCTGAATCTCAACAGGA	1260
Db	11200	TGCAAGTTTCAACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCTGAATCTCAACAGGA	11259
Oy	1261	CTCAGCAACACGGGATCCCACTGTCTGCAAGGACAGCCTGTGACACTCTCTTTCAACCTT	1320
Db	11260	CTCAGCAACACGGGATCCCACTGTCTGCAAGGACAGCCTGTGACACTCTCTTTCAACACCTT	11319
Oy	1321	CATTCTTTCCAGAGATTTGAGAAATGTTATCTCTCCAGGCCCTGACACCCATGTCCTCT	1380
Db	11320	CATTCTTTCCAGAGATTTGAGAAATGTTATCTCTCTCCAGGCCCTGACACCCATGTCCTCT	11379
Oy	1381	GGACTCAGGGTCTGTCTTCCCAACATTGGGCTGACCGTGTCTCTTAGTTGAACCTCTGGG	1440
Db	11380	GGACTCAGGGTCTGTCTTCCCAACATTGGGCTGACCGTGTCTCTTAGTTGAACCTCTGGG	11439
Oy	1441	AACAAATTTCCAAAATGTCTCAAGGCGGGGGTGTGCTCAATCTCTCTGGGCACTTTCA	1500

Db	11440	AAACAATTTCCAAAACACTGTCACAGGGCGGGGGTTCGCGTCAATCTCCCTGGGGACATTTC	11489
Qy	1501	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAAT	1560
Db	11500	TCCTCAAGCTCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCCAATTTAGTCCAGAAAT	11558
Qy	1561	AAACTGGAAG 1571	
Db	11560	AAACTGGAAG 11570	
RESULT 3			
LOCUS	CO874961	11570 bp	DNA linear PAT 27-SEP-2004
DEFINITION	Sequence 5 from Patent WO2004075713.		
ACCESSION	CO874961		
VERSION	CO874961.1	GI:52748060	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1	Diamandis, E. P.	
AUTHORS	Multiple marker assay for detection of ovarian cancer		
TITLE	Patent: WO 2004075713-A 5 10-SEP-2004;		
JOURNAL	MOUNT SINAI HOSPITAL CORPORATION (CA)		
FEATURES	Location/Qualifiers		
SOURCE	1. 11570		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query Match	100.0%;	Score 1571;	DB 6; Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1571;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	AGGAGGATATGGGGAAATTGAAGACAGAGAAACAAATTAATGCCAAGCAATGATTTCTA	60
Db	10000	AGGAGGATATGGGGAAATTGAAGACAGAGAAACAAATTAATGATGCCAAGCAATGATTTCTA	10058
Qy	61	TTGGGAGTGAATTTGCCCTTAGAAGACACTGGCAATACAGAGACATTTTGGTTGTCA	120
Db	10060	TTGGGAGTGAATTTGCCCTTAGAAGACACTGGCAATACAGAGACATTTTGGTTGTCA	10119
Qy	121	CAACTAATAGAGGGGCAATTAATCGCAATTAATGATAGATGCCAAGTGTGCTTCAAC	180
Db	10120	CAACTAATATGAGGGGCAATTAATCGCAATTAATGATAGATGCCAAGTGTGCTTCAAC	10179
Qy	181	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTATCCAGCTTCAGATGCCAC	240
Db	10180	ATGCTATGATGCACACGGCAGGCTCCACAAACCAATTATCCAGCTTCAGATGCCAC	10239
Qy	241	AGTCCCAAGTCAGGAACCCCTCATCCAGGGCTGGAACCGTATTTTTCAGAAAGGAG	300
Db	10240	AGTCCCAAGTCAGGAACCCCTCATCCAGGGCTGGAACCGTATTTTTCAGAAAGGAG	10289
Qy	301	GTTATAGAGTGGGTGTGTGGAGAAATGGGGAAGGAAGTGTGTCTCAAGTAAGAAATTA	360
Db	10300	GTTATAGAGTGGGTGTGTGGAGAAATGGGGAAGGAAGTGTGTCTCAAGTAAGAAATTA	10359
Qy	361	GGCTTGACAGGCTGAGGGGGAAGTGAAGAGAAAGGAGCGGAGAGATCAAGATGA	420
Db	10360	GGCTTGACAGGCTGAGGGGGAAGTGAAGAGAAAGGAGCGGAGAGATCAAGATGA	10419
Qy	421	GGGAGACAGGCTGGAACAGAAATGAGACGAAGATTGAGATGTGAGAGGAAGGCTCA	480
Db	10420	GGGAGACAGGCTGGAACAGAAATGAGACGAAGATTGAGATGTGAGAGGAAGGCTCA	10479
Qy	481	CAGACCCCCCAAAATGATGTGTGACCAACAGGAATCTTGAAAGAGAAAGATGGAATGAG	540
Db	10480	CAGACCCCCCAAAATGATGTGTGACCAACAGGAATCTTGAAAGAGAAAGATGGAATGAG	10539

```
QY 541 AGTGA CAAATGGGGTCTAAAGTTGA ACTTGAGGCGCAGGCATGTGCTCAGCCTGTA 600
Db 10540 AGTGA CAAATGGGGTCTAAAGTTGA ACTTGAGGCGCAGGCATGTGCTCAGCCTGTA 10599
QY 601 ATCCCAACACTTTGAGGCGCTGAGGTGGGGAATCACTTGAAGCCAGAGATTGCAAGCCAG 660
Db 10600 ATCCCAACACTTTGAGGCGCTGAGGTGGGGAATCACTTGAAGCCAGAGATTGCAAGCCAG 10659
QY 661 CTTGGCCCAATGTTGAAAACCCGCTCTACAAAAAATATACAAAAATTAAGCCGGGTCT 720
Db 10660 CTTGGCCCAATGTTGAAAACCCGCTCTACAAAAAATATACAAAAATTAAGCCGGGTCT 10719
QY 721 GGTGATGACACCTGTAGTCA CAGCTACTTGGAGGCTGAGGAGGAGAGATTGCTTGAAC 780
Db 10720 GGTGATGACACCTGTAGTCA CAGCTACTTGGAGGCTGAGGAGGAGAGATTGCTTGAAC 10779
QY 781 CCGGAGATGAGAGGCTGAGTGAAGCTGAGGTCAGGCCACTGCGCTCCAACTGGGCAACA 840
Db 10780 CCGGAGATGAGAGGCTGAGTGAAGCTGAGGTCAGGCCACTGCGCTCCAACTGGGCAACA 10839
QY 841 GAGTAGACTCCATCTCAAAAAAAGCTGGAATTTGAGTGAATTAATTAACATT 900
Db 10840 GAGTAGACTCCATCTCAAAAAAAGCTGGAATTTGAGTGAATTAATTAACATT 10899
QY 901 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Db 10900 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10959
QY 961 CTGTATCTTCCATCTCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 10960 CTGTATCTTCCATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11019
QY 1021 GGTCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db 11020 GGTCTCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11079
QY 1081 TCTCTCTCTGAGGATGATCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 11080 TCTCTCTCTGAGGATGATCTGAGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11139
QY 1141 GTGTCTGTGGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 11140 GTGTCTGTGGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11199
QY 1201 TCCAAATTGACCAAGTGAATCCAGAAACATTCAGAGCCAACTCTGAGTCAATCCAGGA 1260
Db 11200 TCCAAATTGACCAAGTGAATCCAGAAACATTCAGAGCCAACTCTGAGTCAATCCAGGA 11259
QY 1261 CTCAGACACCGGCATCCCACTGTGTGAGGAGCAGCCCTGACACTCTTTCAGACCTT 1320
Db 11260 CTCAGACACCGGCATCCCACTGTGTGAGGAGCAGCCCTGACACTCTTTCAGACCTT 11319
QY 1321 CATTCCTTCCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACACCCCAATGTCCT 1380
Db 11320 CATTCCTTCCAGAGATGTTGAGAAATGTTCAATCTCTCCAGCCCTGACACCCCAATGTCCT 11379
QY 1381 GGAAGTCAAGGATGCTGCTTCCCAATGAGGCTGACCGGTGCTCTAGTTGAACCTTGAG 1440
Db 11380 GGAAGTCAAGGATGCTGCTTCCCAATGAGGCTGACCGGTGCTCTAGTTGAACCTTGAG 11439
QY 1441 AACAAATTTCAAAAATGTC CAGGCGGAGGATGCGTCTCAATCTCCCTGGGCACTTTTCA 1500
Db 11440 AACAAATTTCAAAAATGTC CAGGCGGAGGATGCGTCTCAATCTCCCTGGGCACTTTTCA 11499
QY 1501 TCTCTCAAGTCAAGGCGCCATCTCTTCTGTGAGCTTGTGACCAAAATTAATGTCAGAAAT 1560
Db 11500 TCTCTCAAGTCAAGGCGCCATCTCTTCTGTGAGCTTGTGACCAAAATTAATGTCAGAAAT 11559
QY 1561 AAAGTGAAG 1571
Db 11560 AAAGTGAAG 11570
```

```
RESULT 4
AF135028 11570 bp DNA linear PRI 27-JUN-2000
LOCUS Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
DEFINITION AF135028
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Yousef, G.M. and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
JOURNAL J. Biol. Chem. 274 (53), 37511-37516 (1999)
MEDLINE 20076408
PUBMED 10608802
REFERENCE
AUTHORS Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magliara, A. and Obiezu, C.V.
TITLE The new human kallikrein gene family: implications in
carcinogenesis
JOURNAL Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
MEDLINE 21121728
PUBMED 10675891
REFERENCE
AUTHORS Yousef, G.M., Luo, L.Y. and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
source 1. 11570
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101..2131,2210..2293,4762..5023,5763..6019,
6105..6238,11092..11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221..2293,4762..5023,5763..6019,6105..6238,
11092..11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
/protein_id="AAD26429.1"
/db_xref="GI:4589283"
/translation="MATARPMMWVLCALITALLIGTEHYLANNDVSCDHSNTPS
GSNODLAGAGEDARSDDSSRIINGSDCMHTQWQALILRPOLYCGAVLHPQW
LUTAHCKKVFRLVGLHYSLSPYSESQOMFOGVKSIPHPGYSHRGSNDLMLIKN
RRIPTKQVRDINVSNSHCSAGTCLVSGKVTSSPOYHPKVLQCLNISVLSORCE
DAYRQIDTPMPCADNRKGRDSCGDSGPPVNCGLGVLVSGDYPCARRPRRBYVT
NLCKFTKIQETIQANS"
ORIGIN
Query Match 100.0%; Score 1571; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AGGAGGATATGGGGAATGAAGACAGAAACAAATTAATGTCAGGAAATGATTTCTA 60
Db 10000 AGGAGGATATGGGGAATGAAGACAGAAACAAATTAATGTCAGGAAATGATTTCTA 10059
QY 61 TTGGAGATGATTTCTGCCCCCTAGAAAGACCTGGCAATACAGAGACATTTTGTGTGCA 120
Db 10060 TTGGAGATGATTTCTGCCCCCTAGAAAGACCTGGCAATACAGAGACATTTTGTGTGCA 10119
QY 121 CAACATATGAGGGGGGCACTTACTGGCACTATGATGATGATGATGATGATGATGATGATGAT 180
Db 10120 CAACATATGAGGGGGGCACTTACTGGCACTATGATGATGATGATGATGATGATGATGAT 10179
QY 181 ATGCTATGATGACACGCGAGGCTTCCAAACCAATTAATCCAGCTTCAGATGCCAC 240
```

```

Db 10180 ATGTATGATGACACAGGAGGCTCCACAACAAATATTCAGCTTACAGATGCCAC 10239
Qy 241 AGTCCCAAGATTCAGAGAACTCTATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 300
Db 10240 AGTCCCAAGATTCAGAGAACTCTATCCAGGGGCTGAGAACCGTATTTTTCAGAAAGGAG 10299
Qy 301 GTATTAAGATGGGTGGTGGTGAATGGGGAAGGAAGGTGTGTCTCAATAGAGAAATTA 360
Db 10300 GTATTAAGATGGGTGGTGGTGAATGGGGAAGGAAGGTGTGTCTCAATAGAGAAATTA 10359
Qy 361 GGCCTGACAGAGCTGAGAGGGAGAGTGAAGAGAAAGGAGCCGAGAGATACAGATGA 420
Db 10360 GGCCTGACAGAGCTGAGAGGGAGAGTGAAGAGAAAGGAGCCGAGAGATACAGATGA 10419
Qy 421 GGGAGACAGGCTGGAACAGAAAGTGAAGACAGAGATTCAGATGTGAGAGAAAGGCTCA 480
Db 10420 GGGAGACAGGCTGGAACAGAAAGTGAAGACAGAGATTCAGATGTGAGAGAAAGGCTCA 10479
Qy 481 CAGACCCCGCGAAATGATGTGTGAGACAGAGATCTGGAAGAGAAAGATGAGTGGAG 540
Db 10480 CAGACCCCGCGAAATGATGTGTGAGACAGAGATCTGGAAGAGAAAGATGAGTGGAG 10539
Qy 541 AGTGAACAAATGGGCTTAAAGTTGAACTTGAAGCCAGAGCATGTGCTCACGCTGTGA 600
Db 10540 AGTGAACAAATGGGCTTAAAGTTGAACTTGAAGCCAGAGCATGTGCTCACGCTGTGA 10599
Qy 601 ATCCCAACACTTTGAGAGGCTGAGAGGTGGCGCAATCACTTGAAGCCAGAGATTCAGACAG 660
Db 10600 ATCCCAACACTTTGAGAGGCTGAGAGGTGGCGCAATCACTTGAAGCCAGAGATTCAGACAG 10659
Qy 661 CTGGGCAACATGGTGAACCCCGCTCTCAAAAAAATTAACAAATTAACCGGGTGT 720
Db 10660 CTGGGCAACATGGTGAACCCCGCTCTCAAAAAAATTAACAAATTAACCGGGTGT 10719
Qy 721 GGTGATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGACAGAGAAATTCGTTGAC 780
Db 10720 GGTGATGAGACACTGTAGTCAAGCTACTTGGAGGCTGAGGACAGAGAAATTCGTTGAC 10779
Qy 781 CCGGAGATGAGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 840
Db 10780 CCGGAGATGAGGCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 10839
Qy 841 GAGTAACACTCCATCTCAAAAAAAGCTGATTTGAGTAATTAATAAATTAATAAATTAATAAATTA 900
Db 10840 GAGTAACACTCCATCTCAAAAAAAGCTGATTTGAGTAATTAATAAATTAATAAATTAATAAATTA 10899
Qy 901 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 960
Db 10900 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10959
Qy 961 CTGTACTTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
Db 10960 CTGTACTTTCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11019
Qy 1021 GGTCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Db 11020 GGTCTCTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11079
Qy 1081 TCTCTCTCTGAGAGGATTCAGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Db 11080 TCTCTCTCTGAGAGGATTCAGAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11139
Qy 1141 GTGTCTGTGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Db 11140 GTGTCTGTGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 11199
Qy 1201 TGCAGATTCACCAAGTGAATTCAGAGAAACATTCAGAGCCAACTCTCTGAGTCAATCCAGAGA 1260
Db 11200 TGCAGATTCACCAAGTGAATTCAGAGAAACATTCAGAGCCAACTCTCTGAGTCAATCCAGAGA 11259
Qy 1261 CTCAGACACCGGATCCCACTGCTGACAGGAGACAGCCCTGACACTCTTTGAGACCTT 1320

```

```

Db 11260 CTCAGACACCGGATCCCACTGCTGACAGGACAGCCCTGACACTCTTTGAGACCTT 11319
Qy 1321 CATTCCTCCAGAGATGTGAATGTTCATCTCCAGCCCTGACCCCATGTCTCT 1380
Db 11320 CATTCCTCCAGAGATGTGAATGTTCATCTCCAGCCCTGACCCCATGTCTCT 11379
Qy 1381 GGACTCAGAGTCTGTCTTCCCAATTTGGGCTGACCGTGTCTCTAGTTGAACCTCTGG 1440
Db 11380 GGACTCAGAGTCTGTCTTCCCAATTTGGGCTGACCGTGTCTCTAGTTGAACCTCTGG 11439
Qy 1441 AACAAATTCACAAATCTGTCCAGGGCGGGGCTGTCTCAATCTCCCTGGGCACTTTCA 1500
Db 11440 AACAAATTCACAAATCTGTCCAGGGCGGGGCTGTCTCAATCTCCCTGGGCACTTTCA 11499
Qy 1501 TCCTCAAGCTCAGGGCCCAATCCCTCTCTGACCTGACCCAAATTTAGTCCAGAAAT 1560
Db 11500 TCCTCAAGCTCAGGGCCCAATCCCTCTCTGACCTGACCCAAATTTAGTCCAGAAAT 11559
Qy 1561 AAACGTAGAG 1571
Db 11560 AAACGTAGAG 11570

RESULT 5
AC011483/c 107487 bp DNA linear PRI 01-JUL-2002
LOCUS AC011483
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.
ACCESSION AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submision
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
1. 107487
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C22"
/64998..65494
/note="NOTE: Shatter libraries failed to resolve
dinucleotide repeat. Unsure number of repeat copies
64998-65494. Forced join 65015."

Origin
Query Match 100.0%; Score 1571; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


CDS	<p>/gene="KLK1" /product="renal kallikrein" complement(join(6963..7118,7668..7804,7923..8212,9482..9641,11472..11517)) /gene="KLK1" /note="serine protease" /codon_start=1 /product="renal kallikrein" /protein_id="AA033353.1" /db_xref="GI:11244758" /translation="MMFLVLCALSLSGTGAPPIQSRIYGVGECRHSQPMQVALYHFSYFQCGGILVHROVWLTAAHCISDNQYLMGLHNFEDENTAOPIVSSPFGPNMSLEHNRADSDYSHDMLRLSEPADITDAVKVLEPTREEVSGCLASGMSIEBENFSPPDLOCVLKITLBNCKKAHOKXTDPLCVHLEGGKDTVDSGGLMCDGVLOGVTSKRYPCGTIPKPSVANRVLSYKNIETIAENS"</p>
mRNA	<p>complement(join(<13552..13704,14377..14530,14678..14917,15416..15569,19204..19246)) /product="ACO protease" complement(join(13552..13704,14377..14530,14678..14917,15416..15569,19204..19246)) /note="serine protease" /codon_start=1 /product="ACO protease" /protein_id="AA033354.1" /db_xref="GI:11244759" /translation="MMVLTLSPLASTAODGKLEGEDECAPHSGPMQVALYERGRFNGCASLSPHWLSAHCOSRFRVRLGERHNRKRDPEQLTTSRVTFPHRYEARSHRNDMLRLVOPARLNPQVPAPVLPFRCPHPEACVSGMGLVPLSSPPLDTHLCANISLISDSCKDSYPGRLTNTVWCAEGRGSCEGDSGGLVCGGIIQGIIVSGWDVPCDNTTKPGVYKVCYHEWIRETMKRN"</p>
gene	<p><42595..>47769 /gene="KLK3" join(<42595..42640,43880..44039,45669..45955,46099..46235,47614..>47769) /gene="KLK3" /product="prostate specific antigen" join(<42595..42640,43880..44039,45669..45955,46099..46235,47614..47769) /gene="KLK3" /note="serine protease" /codon_start=1 /product="prostate specific antigen" /protein_id="AA033355.1" /db_xref="GI:11244760" /translation="MMVYVFLTSLVTWIGAPLILSRIYGVGECRHSQPMQVALYHRSRAVCGGVVHPQWVLTAAHCIRNKSQVILGHRSLFHPEDTQVQVSHSPHPLYDMSLKKRFLRPBDDSSHDLMLRLSEBAELTDKVKWDLPTQEPALGTTCYASGMSIEPEEFLPKLQCVLDLHVISNDYCAQVHPQKTKFPMLCAGRWGKSTCSGDSGGLVCNGVLOGITSMGSEPCALPERPSLYTKVHYRKMIQDTIANP"</p>
mRNA	<p><61139..>66229 /gene="KLK2" join(<61139..61184,62391..62550,64142..64428,64542..64678,66074..>66229) /gene="KLK2" /product="glutular kallikrein" join(<61139..61184,62391..62550,64142..64428,64542..64678,66074..66229) /gene="KLK2" /note="serine protease" /codon_start=1 /product="glutular kallikrein" /protein_id="AA033356.1" /db_xref="GI:11244761" /translation="MMDLVLSIALSVGCTGAVPLIOSRIYGVGECRHSQPMQVALYHSGAHCGGVVHPQWVLTAAHCIRNKSQVILGHRSLFHPEDTQVQVSHSPHPLYNMSLKKRFLRPBDDSSHDLMLRLSEBAELTDKVKWDLPTQEPALGTTCYASGMSIEPEEFLPKLQCVLDLHVISNDYCAQVHPQKTKFPMLCAGRWGKSTCSGDSGGLVCNGVLOGITSMGSEPCALPERPAVYTKVHYRKMIQDTIANP"</p>
CDS	<p>complement(join(<94602..94755,96027..96163,96247..96497,96919..97081,98345..>98405)) /product="serine protease" complement(join(94602..94755,96027..96163,96247..96497,96919..97081,98345..98405)) /note="serine protease; also called ovaasin" /codon_start=1 /product="neutropain" /protein_id="AA033357.1" /db_xref="GI:11244762"</p>
gene	<p>complement(<131301..>140303) /gene="KLK5" /note="synonym: SCTE" complement(join(<131301..131456,136310..136443,136529..136785,137525..137690,140255..>140303)) /gene="KLK5" /product="stratum corneum trypsin-like serine protease" complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..140303)) /gene="KLK5" /codon_start=1 /product="stratum corneum trypsin-like serine protease" /protein_id="AA033358.1" /db_xref="GI:11244763" /translation="MMVLCALITALLGYTDARSDSSRIINGSDCMHTQPMOALLRPNLYCGAVLVHPQWLTAAHCIRNKSQVILGHRSLFHPEDTQVQVSHSPHPLYNMSLKKRFLRPBDDSSHDLMLRLSEBAELTDKVKWDLPTQEPALGTTCYASGMSIEPEEFLPKLQCVLDLHVISNDYCAQVHPQKTKFPMLCAGRWGKSTCSGDSGGLVCNGVLOGITSMGSEPCALPERPAVYTKVHYRKMIQDTIANP"</p>
mRNA	<p>complement(join(<146834..146986,149628..149764,151186..151433,155052..155208,155948..155987)) /product="protease M" complement(join(146834..146986,149628..149764,151186..151433,155052..155208,155948..155987)) /note="serine protease; also called neutrobin or zyme" /codon_start=1 /product="protease M" /protein_id="AA033359.1" /db_xref="GI:11244764" /translation="MKKMLVYSLTAAAEQNKLVHSGPCKTSHPYOALYTSCHLTCGGVLIHPVLTAAHCIRNKSQVILGHRSLFHPEDTQVQVSHSPHPLYNMSLKKRFLRPBDDSSHDLMLRLSEBAELTDKVKWDLPTQEPALGTTCYASGMSIEPEEFLPKLQCVLDLHVISNDYCAQVHPQKTKFPMLCAGRWGKSTCSGDSGGLVCNGVLOGITSMGSEPCALPERPAVYTKVHYRKMIQDTIANP"</p>
CDS	<p>complement(<165420..165575,167672..167808,168124..168371,169651..169798,170211..>170283)) /gene="KLK7" /product="stratum corneum chymotryptic enzyme" complement(join(165420..165575,167672..167808,168124..168371,169651..169798,170211..170283)) /gene="KLK7" /note="serine protease" /codon_start=1 /product="stratum corneum chymotryptic enzyme" /protein_id="AA033360.1" /db_xref="GI:11244765" /translation="MARSLLPLQILLSLAETAGEBAQDKIIDGAPCANGSHPMQVALLSGNQLHCGGVVHPQWLTAAHCIRNKSQVILGHRSLFHPEDTQVQVSHSPHPLYNMSLKKRFLRPBDDSSHDLMLRLSEBAELTDKVKWDLPTQEPALGTTCYASGMSIEPEEFLPKLQCVLDLHVISNDYCAQVHPQKTKFPMLCAGRWGKSTCSGDSGGLVCNGVLOGITSMGSEPCALPERPAVYTKVHYRKMIQDTIANP"</p>
gene	<p>complement(join(<183943..184098,185635..185768,187865..188127,188293..188452,188967..>189036)) /product="neutropain" complement(join(183943..184098,185635..185768,187865..188127,188293..188452,188967..189036)) /note="serine protease; also called ovaasin" /codon_start=1 /product="neutropain" /protein_id="AA033361.1" /db_xref="GI:11244766"</p>
mRNA	<p>complement(join(<183943..184098,185635..185768,187865..188127,188293..188452,188967..189036)) /note="serine protease; also called ovaasin" /codon_start=1 /product="neutropain" /protein_id="AA033361.1" /db_xref="GI:11244766"</p>
CDS	<p>complement(join(<183943..184098,185635..185768,187865..188127,188293..188452,188967..189036)) /note="serine protease; also called ovaasin" /codon_start=1 /product="neutropain" /protein_id="AA033361.1" /db_xref="GI:11244766"</p>

/translation="MGRPRRAKTMFLLLIGAMWAGHRAOEDKVLGHCOPHSQ
PWOALFOGQOLLCGVLVGVGMVLAALCKKPKYLRVLAQKQDPEBEPVVO
SLPHCYNSDVEDHNDHMLLORBOALSKKXKPISTADCTOPKOCSTYSGMGV
TSPRNPPTLNCARVKEIPROKCEDAYVGOITDQVNCSSKSGADTCQDGBGFLVC
DCALOGITSMGSDPCGRBSPRYVYINICYLDMIKKIIJSSKG.
complement (join<190980, .191129,191573, .191709,
194324, .194589,197048, .197204,197370, .1974121)
/product="Kallikrein-like 3"
complement (join190980, .191129,191573, .191709,
194324, .194589,197048, .197204,197370, .1974121)
/note="serine protease"
/codon_start=1
/product="Kallikrein-like 3"
/protein_id="AA033362.1"
/db_xref="GI:11244767"
/translation="MKGLCALSLSLAGWADTRATIGAEBCRPNOSPMOGLPHLT
RLFCATLISDRLTLAARCRPYLMVLRGHHMLKMBEPELFRVTFPRPRKRD
LSANDHNDIMLRIPROKRLSPAVOPLSOTCVSPKQCLISGVAVSPKALFV
TLQCANISLEKLCHEMAYPGHISDMLCAGMEGRSGSQDSDGPLVNCNGLTAGV
SGAEPSCRPRPRPAVYTSVCHYLDWIQEIEMEN"

Query Match 100.0%; Score 1571; DB 9; Length 230000;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGGATATGGGATTTAGAGACAGAAACAAATTAGTCCAGAGCATTTCTA 60
DB 132548 AGGAGGATATGGGATTTAGAGACAGAAACAAATTAGTCCAGAGCATTTCTA 132489
QY 61 TTGGAGATATTTCTGCCCCCTAGAAACATCGGCATATCAAGAGACATTTTGGTTGCA 120
DB 132488 TTGGAGATATTTCTGCCCCCTAGAAACATCGGCATATCAAGAGACATTTTGGTTGCA 132429
QY 121 CAATATATAGAGGGGATTTACTGCGCACTAATGATAGTCAAGTGTGCTTCAAC 180
DB 132428 CAATATATAGAGGGGATTTACTGCGCACTAATGATAGTCAAGTGTGCTTCAAC 132369
QY 181 ATGCTATATGACACCGGAGGCTTCAACAAACCATTTATTCAGATGCCAC 240
DB 132368 ATGCTATATGACACCGGAGGCTTCAACAAACCATTTATTCAGATGCCAC 132309
QY 241 AGTGCCAGATGAGGAACCTCTATTCAGAGGCTGAGAACCGTATTTTTCAGAAAGGAG 300
DB 132308 AGTGCCAGATGAGGAACCTCTATTCAGAGGCTGAGAACCGTATTTTTCAGAAAGGAG 132249
QY 301 GTATAGATATGGGATTTGTTGAGAAATGGGAGAAAGTGTGTCCAGTAAGAAATTA 360
DB 132248 GTATAGATATGGGATTTGTTGAGAAATGGGAGAAAGTGTGTCCAGTAAGAAATTA 132189
QY 361 GGCCTTGACAGGCTGAGAGGAGAGTGAAGAGAAAGGAGGCGAGAGATACAGATGA 420
DB 132188 GGCCTTGACAGGCTGAGAGGAGAGTGAAGAGAAAGGAGGCGAGAGATACAGATGA 132129
QY 421 GGGAGACAGGCTGAGAACAGAAAGTGAACAGAAATTCAGATGTTGAGAGAGAGGTTCA 480
DB 132128 GGGAGACAGGCTGAGAACAGAAAGTGAACAGAAATTCAGATGTTGAGAGAGAGGTTCA 132069
QY 481 CAGACCCCCCGGAAATGATGTGTGAACAAGAAATCTGAAAGAGAAATGAGATGAGAG 540
DB 132068 CAGACCCCCCGGAAATGATGTGTGAACAAGAAATCTGAAAGAGAAATGAGATGAGAG 132009
QY 541 AGTGAACAAATGGGCTTAAAGTTGAACCTTGAGGCGAGAGCATGTGCTCAGCCTGTA 600
DB 132008 AGTGAACAAATGGGCTTAAAGTTGAACCTTGAGGCGAGAGCATGTGCTCAGCCTGTA 131949
QY 601 ATCCCAACACTTTTGAAGCTGAGGTGGCGGAATCACTTGAGGCGAGAGTTTGAAGCCAG 660
DB 131948 ATCCCAACACTTTTGAAGCTGAGGTGGCGGAATCACTTGAGGCGAGAGTTTGAAGCCAG 131889
QY 661 CTTGGCAACATGTGAAACCCCGTCTCAACAAAAAATTAACAAAAATTAACCCGGGTCT 720
DB 131888 CTTGGCAACATGTGAAACCCCGTCTCAACAAAAAATTAACAAAAATTAACCCGGGTCT 131829

QY 721 GGTGATGACACCTGTAGTCAAGCTACTTGGAGGCTGAGGCGAGAGAAATTTGTAAC 780
DB 131828 GGTGATGACACCTGTAGTCAAGCTACTTGGAGGCTGAGGCGAGAGAAATTTGTAAC 131769
QY 781 CCGGAGATGAGGCTGAGTGAAGTGAAGTCAAGGCTGAGGCTTCAACTTGGGACACA 840
DB 131768 CCGGAGATGAGGCTGAGTGAAGTGAAGTCAAGGCTGAGGCTTCAACTTGGGACACA 131709
QY 841 GAGTAAGACTCATCTCAAAAAAAGAGTGAATTTGAGTGAATTAATTAACATT 900
DB 131708 GAGTAAGACTCATCTCAAAAAAAGAGTGAATTTGAGTGAATTAATTAACATT 131649
QY 901 CTCCTCTCTCTCTCTTGGCTGTGTCTGCATCTGTCTTCTTTCGATTTCTTCATCT 960
DB 131648 CTCCTCTCTCTCTCTTGGCTGTGTCTGCATCTGTCTTCTTTCGATTTCTTCATCT 131589
QY 961 CTGTACTTTTCATCTCTGTGTGTCTGTTCCTCATCTGCTTTCATCTATGAGCATCTGTG 1020
DB 131588 CTGTACTTTTCATCTCTGTGTGTCTGTTCCTCATCTGCTTTCATCTATGAGCATCTGTG 131529
QY 1021 GGTCTCTCATGTCTCTTCTGCGCCACTTGGCAGACATCTGCTCTCATGCCCCCTT 1080
DB 131528 GGTCTCTCATGTCTCTTCTGCGCCACTTGGCAGACATCTGCTCTCATGCCCCCTT 131469
QY 1081 TCTCTCTGACAGGATTTCTGGGGGAGCTGTGTCTGCAATGAGTCTCCCTGAGGAGCTC 1140
DB 131468 TCTCTCTGACAGGATTTCTGGGGGAGCTGTGTCTGCAATGAGTCTCCCTGAGGAGCTC 131409
QY 1141 GTGTCTTGGGAGATTAACCTTGTGTGCTGCGGCCCAACAGACGGGAGTCTACACGAACCTC 1200
DB 131408 GTGTCTTGGGAGATTAACCTTGTGTGCTGCGGCCCAACAGACGGGAGTCTACACGAACCTC 131349
QY 1201 TGCAGTTCAACCAATGTGATCCAGGAAACATCCAGGCAACTCTGATGATCCAGAGA 1260
DB 131348 TGCAGTTCAACCAATGTGATCCAGGAAACATCCAGGCAACTCTGATGATCCAGAGA 131289
QY 1261 CTCAGCACACCGGACATCCCACTGTGCTGACAGGAGACAGCTCTGCTTTCAGACCT 1320
DB 131288 CTCAGCACACCGGACATCCCACTGTGCTGACAGGAGACAGCTCTGCTTTCAGACCT 131229
QY 1321 CATTTCTTCCAGAGATTTGAGAAATTTCAATCTCTCAGGCTTCAACCCCATGCTCTCT 1380
DB 131228 CATTTCTTCCAGAGATTTGAGAAATTTCAATCTCTCAGGCTTCAACCCCATGCTCTCT 131169
QY 1381 GGACTCAGGAGTGTGTTCCCAACATGAGGCTGACGAGTCTCTAGTTGAACCTGAG 1440
DB 131168 GGACTCAGGAGTGTGTTCCCAACATGAGGCTGACGAGTCTCTAGTTGAACCTGAG 131109
QY 1441 AACAAATTTCAAAACCTGTCAAGGCGGAGGTTGCTCAATCTCCCTGAGGCACTTTCA 1500
DB 131108 AACAAATTTCAAAACCTGTCAAGGCGGAGGTTGCTCAATCTCCCTGAGGCACTTTCA 131049
QY 1501 TCTCTAAGCTCAGGGCCCATCTCTCTCTGCAAGCTCTGACCCCAATTTTATGTTCCGAAT 1560
DB 131048 TCTCTAAGCTCAGGGCCCATCTCTCTCTGCAAGCTCTGACCCCAATTTTATGTTCCGAAT 130989
QY 1561 AAACCTGAGAAG 1571
DB 130988 AAACCTGAGAAG 130978

RESULT 7
AC027602/c
LOCUS
DEFINITION
SEQUENCE, 12 unordered pieces.
ACCESSION
AC027602
VERSION
AC027602.4 GI:11178143
KEYWORDS
HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
AUTHORS 1 (bases 1 to 217346)
TITLE Birren,B., Linton,L., Nuebaum,C. and Lander,E.
JOURNAL Homo sapiens chromosome 19, clone RP11-795B6
REFERENCE
AUTHORS 2 (bases 1 to 217346)
TITLE Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Bogulavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choegl,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,D.S.,
Dodg,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Kleith,J., Larocque,K., Lamazeres,R., Landers,T., Lehocsky,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McKean,P., McGuck,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisanti,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfaye,S., Theodore,J., Titrill,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE
JOURNAL Direct Submission
AUTHORS Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 217346)

TITLE
JOURNAL Direct Submission
AUTHORS Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Nov 15, 2000 this sequence version replaced gi:1136831.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center Project name: L9166
Center clone name: 795_B-6
----- Summary Statistics
Sequencing vector: M13; M77815; 31% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 213655 bases at least Q30
Consensus quality: 215056 bases at least Q20

Insert size: 194000; agarose-fp
Insert size: 216246; sum-of-contigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.
NOTE: This is a 'working draft' sequence. It currently
consists of 12 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 8149: contig of 8149 bp in length
8150 8249: gap of 100 bp
8250 9592: contig of 1343 bp in length
9593 9692: gap of 100 bp
9693 10733: contig of 1041 bp in length
10734 10833: gap of 100 bp
10834 13519: contig of 2686 bp in length
13520 13619: gap of 100 bp
13620 17510: contig of 3891 bp in length
17511 17610: gap of 100 bp
17611 24602: contig of 6992 bp in length
24603 24702: gap of 100 bp
24703 35434: contig of 10732 bp in length
35435 35534: gap of 100 bp
35535 124474: contig of 88940 bp in length
124475 124574: gap of 100 bp
124575 134664: contig of 10090 bp in length
134665 134764: gap of 100 bp
134765 162343: contig of 27579 bp in length
162344 162443: gap of 100 bp
162444 208917: contig of 46474 bp in length
208918 209018: gap of 100 bp
209018 217346: contig of 8329 bp in length.

FEATURES

source 1. 217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_1lb="RPCT-11 Human Male BAC"
1. 8149
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
8250..9592
/note="assembly_fragment"
9693..10733
/note="assembly_fragment"
10834..13519
/note="assembly_fragment"
13620..17510
/note="assembly_fragment"
17611..24602
/note="assembly_fragment"
24703..35434
/note="assembly_fragment"
35535..124474
/note="assembly_fragment"
124575..134664
/note="assembly_fragment"
134765..162343
/note="assembly_fragment"
162444..208917
/note="assembly_fragment"
209018..217346
clone_end:T7
vector_side:right"

ORIGIN

Query Match	99.0%; Score 1555.2; DB 2; Length 217346;	
Best Local Similarity	99.7%; Pred. No. 0;	
Matches 1568; Conservative	0; Mismatches 3; Indels 1; Gaps 1;	
QY	1 AGGAGATATGGGAAATTTGAAGACAGAAACAAATTTAGTCCAGCCGAATGGATTCTGA	60
DB	194423 AGGAGATATGGGAAATTTGAAGACAGAAACAAATTTAGTCCAGCCGAATGGATTCTGA	194364
QY	61 TTGGAGATATTCCTGCCCCCTAGAAAGACATCGCAATACAGAGAGACATTTTGGTTGTA	120
DB	194363 TTGGAGATATTCCTGCCCCCTAGAAAGACATCGCAATACAGAGAGACATTTTGGTTGTA	194304
QY	121 CAATATATGAGGGGCAATTAATGAGCACTTAATGATAGATGCAAGTGTGCTTCAAC	180
DB	194303 CAATATATGAGGGGCAATTAATGAGCACTTAATGATAGATGCAAGTGTGCTTCAAC	194244
QY	181 ATGCTATGATGCAACAGGCGCTTCAACAACAACATTTATCAGCTTCAATGCCAC	240
DB	194243 ATGCTATGATGCAACAGGCGCTTCAACAACAACATTTATCAGCTTCAATGCCAC	194184
QY	241 AGTCCCAATGAGGAGACCTTCATCAGGGGCTGGAACCGATTTTGGAGAGGGAG	300
DB	194183 AGTCCCAATGAGGAGACCTTCATCAGGGGCTGGAACCGATTTTGGAGAGGGAG	194124
QY	301 GTATAAGATGGTGTGAGAGATGGGAGAGAGTGTGTCCAGTAAAGAAATTA	360
DB	194123 GTATAAGATGGTGTGAGAGATGGGAGAGAGTGTGTCCAGTAAAGAAATTA	194064
QY	361 GGCCTGCAACAGCTGAGAGGGGAGAGTGAAGAGAAAGGAGCGAGATACAGATGA	420
DB	194063 GGCCTGCAACAGCTGAGAGGGGAGAGTGAAGAGAAAGGAGCGAGATACAGATGA	194004
QY	421 GGGAGACAGGCTGGAACAGAAAGTGAAGCAAGATTTGAGATGTGAGAGAGAGGCTGA	480
DB	194003 GGGAGACAGGCTGGAACAGAAAGTGAAGCAAGATTTGAGATGTGAGAGAGAGGCTGA	193944
QY	481 CAGACCCCCCAATGATGTGTGAACAAGGAATCTGGAAGAGAGAGATGAGTGAAG	540
DB	193943 CAGACCCCCCAATGATGTGTGAACAAGGAATCTGGAAGAGAGAGATGAGTGAAG	193884
QY	541 AGTGAACAATGGGCTTTAAAGTTGAACCTTGAAGCCAGGCAATGTGGCTCAACGCTGTA	600
DB	193883 AGTGAACAATGGGCTTTAAAGTTGAACCTTGAAGCCAGGCAATGTGGCTCAACGCTGTA	193824
QY	601 ATCCCAACATTTGAGGCTGAGGTGGGGAATCACTTGAAGCCAGAGTGTGAAGCCAG	660
DB	193823 ATCCCAACATTTGAGGCTGAGGTGGGGAATCACTTGAAGCCAGAGTGTGAAGCCAG	193764
QY	661 CCTGGCCAAATGATGTGAACCCCGTCTCTACAAAAAAAATTAACAAATTAACGCGGTGT	720
DB	193763 CCTGGCCAAATGATGTGAACCCCGTCTCTACAAAAAAAATTAACAAATTAACGCGGTGT	193704
QY	721 GGTGATGACACCTGTATGACAGCTACTTGGAGGCTGAAGGAGAGAAATGCTTGAAC	780
DB	193703 GGTGATGACACCTGTATGACAGCTACTTGGAGGCTGAAGGAGAGAAATGCTTGAAC	193644
QY	781 CCGGAGATGAGAGGCTGAGTGAAGCTGAGGCTCACTGGGCTCAACCTGGGGAACA	840
DB	193643 CCGGAGATGAGAGGCTGAGTGAAGCTGAGGCTCACTGGGCTCAACCTGGGGAACA	193584
QY	841 GAGTAAAGATCATCTC-AAAAAAGATGATTTGAGTGAATTAATTAACAT	899
DB	193583 GAGTAAAGATCATCTC-AAAAAAGATGATTTGAGTGAATTAATTAACAT	193524
QY	900 TCTCCCTCTCTCTCTTGGCTGTGTCTCATCTGTCTTTTTCGCAATTTCTTCATC	959
DB	193523 TCTCCCTCTCTCTCTTGGCTGTGTCTCATCTGTCTTTTTCGCAATTTCTTCATC	193464
QY	960 TCTGATCTTTCATCTCTGTGTGTCTTCCATCTGCTTCCATCTAATGGGCAATCTCT	1019
DB	193463 TCTGATCTTTCATCTCTGTGTGTCTTCCATCTGCTTTCATCTAATGGGCAATCTCT	193404

QY	1020 GGGTCTCTCATGTCTCTCTTGTGCCCCCATTTGGCAGATCTGTGCTCTCATGCCCT	1079
DB	193403 GGGTCTCTCATGTCTCTCTTGTGCCCCCATTTGGCAGATCTGTGCTCTCATGCCCT	193344
QY	1080 TTCTCTCTGCGAGGATGATTTCTGGGGGGCTGTGTCTGCAATGGCTCCCTGAGAGACT	1139
DB	193343 TTCTCTCTGCGAGGATGATTTCTGGGGGGCTGTGTCTGCAATGGCTCCCTGAGAGACT	193284
QY	1140 CGTGTCTGGGAGATTAACCTTTGTGCCCGGCAACAGACCGGGTGTCTACAGAACT	1199
DB	193283 CGTGTCTGGGAGATTAACCTTTGTGCCCGGCAACAGACCGGGTGTCTACAGAACT	193224
QY	1200 CTGCAAGTTCAACCAAGTGAATCAGAGAAACATCCAGGCCAACTCTGAGTATCCAGG	1259
DB	193223 CTGCAAGTTCAACCAAGTGAATCAGAGAAACATCCAGGCCAACTCTGAGTATCCAGG	193164
QY	1260 ACTCAGCACACCGGATATCCCACTGTGTGCAAGGACAGCCCTGACATCTTTCAGACC	1319
DB	193163 ACTCAGCACACCGGATATCCCACTGTGTGCAAGGACAGCCCTGACATCTTTCAGACC	193104
QY	1320 TCATTCCTTCCAGAGATGTTGAAGATGTTCACTCTGCAAGCCCTGACCCCATGTCTCC	1379
DB	193103 TCATTCCTTCCAGAGATGTTGAAGATGTTCACTCTGCAAGCCCTGACCCCATGTCTCC	193044
QY	1380 TGAAGTCAAGGCTCTGCTTCCCAATGAGGCTGACCGTGTCTCTAATGAACCTGAG	1439
DB	193043 TGAAGTCAAGGCTCTGCTTCCCAATGAGGCTGACCGTGTCTCTAATGAACCTGAG	192984
QY	1440 GAACAATTTCCAAACTGTCCAGGGGCGGGGGTGTGCTCAATCTCCCTGGGGCACTTTC	1499
DB	192983 GAACAATTTCCAAACTGTCCAGGGGCGGGGGTGTGCTCAATCTCCCTGGGGCACTTTC	192924
QY	1500 ATCTCAAGCTCAGGGGCCATCTCTCTGAGGCTGACCCAAATTTAATGCCAGAAA	1559
DB	192923 ATCTCAAGCTCAGGGGCCATCTCTCTGAGGCTGACCCAAATTTAATGCCAGAAA	192864
QY	1560 TAACTGAGAG 1571	
DB	192863 TAACTGAGAG 192852	
RESULT 8		
AC130782/c		
LOCUS	200792 bp DNA linear HTG 22-NOV-2002	
DEFINITION	Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12	
ORDERED	ordered pieces.	
ACCESSION	AC130782.2 GI:25167101	
VERSION	AC130782.2 HTGS, PHASE2; HTGS, DRAFT.	
KEYWORDS	HTG; HTGS, PHASE2; HTGS, DRAFT.	
SOURCE	Pan troglodytes (chimpanzee)	
ORGANISM	Pan troglodytes	
REFERENCE		
1	(bases 1 to 200792)	
Akhler,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,		
Benjamin,B., Blakeley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,		
Carriaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,		
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,		
Latic,P., Lee-Lin,S.-O., Legaspi,R., Maduro,Q.U., Maduro,V.B.,		
Maruligan,C., Pearson,R., Portnoy,M.E., Prasad,A.,		
Reddik-Duque,N., Schandler,K., Schueler,M.G., Sison,C.,		
Stantirip,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,		
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.		
NISC Comparative Sequencing Initiative		
Unpublished		
JOURNAL		
REFERENCE		
2	(bases 1 to 200792)	
Green,E.D.		
TITLE		
JOURNAL		
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717		
Groveom Circle, Gaithersburg, MD 20877, USA		
3	(bases 1 to 200792)	
Green,E.D.		

TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Government Circle, Gaithersburg, MD 20877, USA
COMMENT On Nov 22, 2002 this sequence version replaced gi:22210452.

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Project name: dhz
Center clone name: 355A20

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data, from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 21500; agarose-fp
Insert size: 199692; sum-of-contigs
Quality coverage: 9.03x in Q20 bases; agarose-fp
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 28306: contig of 28306 bp in length
28307 28406: gap of unknown length
28407 37856: contig of 9450 bp in length
37857 37956: gap of unknown length
37957 73522: contig of 35566 bp in length
73523 73622: gap of unknown length
73623 83567: contig of 9945 bp in length
83568 83667: gap of unknown length
83668 88817: contig of 5150 bp in length
88818 88917: gap of unknown length
88918 125611: contig of 36694 bp in length
125612 125711: gap of unknown length
125712 159879: contig of 34168 bp in length
159880 159979: gap of unknown length
159980 174698: contig of 14719 bp in length
174699 174798: gap of unknown length
174799 186382: contig of 11584 bp in length
186383 186482: gap of unknown length
186483 193344: contig of 6862 bp in length
193345 193444: gap of unknown length
193445 199463: contig of 5919 bp in length
199464 200792: gap of unknown length
200792: contig of 1329 bp in length.

FEATURES
Source
1. .200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"

/db xref="taxon:9598"
/clone="CH251-355A20"
/clone_11b="CH251"
1. .28306
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
/note="assembly_fragment"
37957. .73522
/note="assembly_fragment"
73623. .93567
/note="assembly_fragment"
83668. .88817
/note="assembly_fragment"
88918. .125611
/note="assembly_fragment"
125712. .159879
/note="assembly_fragment"
159980. .174698
/note="assembly_fragment"
174799. .186382
/note="assembly_fragment"
186483. .193344
/note="assembly_fragment"
193445. .199363
/note="assembly_fragment"
199464. .200792
/note="assembly_fragment"
clone_end:17
vector_side:right"

ORIGIN

Query Match 94.3%; Score 1482; DB 2; Length 200792;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1542; Conservative 0; Mismatches 25; Indels 5; Gaps 4;

1 AGGAGAGTATGGGAGTTTGAAGACAGAAACCAATTGTCACAGCAATGATTTCTA 60
69290 AGGAGAGTATGGGAGTTTGAAGACAGAAACCAATTGTCACAGCAATGATTTCTA 69231
QY TTGGAGTAGTATTCGCTCCCTAGACACTGCGAATACAGAGACATTTTGGTTGTC 120
Db TTGGAGTAGTATTCGCTCCCTAGACACTGCGAATACAGAGACATTTTGGTTGTC 69171
QY 61 CAACTATATGAGAGGCGCATTTACTGCGAATATGATAGATGCCAATGCTGTTCAAC 180
Db CAACTATATGAGAGGCGCATTTACTGCGAATATGATAGATGCCAATGCTGTTCAAC 69111
QY 121 ATGCTATGATGACACAGGCGCTTCCACAAACCATTTACGCTTCAATGCCAC 240
Db ATGCTATGATGACACAGGCGCTTCCACAAACCATTTACGCTTCAATGCCAC 69051
QY 69110 GTGCTATGATGACACAGGCGCTTCCACAAACCATTTACGCTTCAATGCCAC 240
Db GTGCTATGATGACACAGGCGCTTCCACAAACCATTTACGCTTCAATGCCAC 69051
QY 241 AGTCCCAATGAGAGAGCCCTCATCCAGAGGCTGGAACCGTATTTTGCAGAGAGAG 300
Db AGTCCCAATGAGAGAGCCCTCATCCAGAGGCTGGAACCGTATTTTGCAGAGAGAG 68991
QY 301 GATTAAGATGAGTGGTGTGAGAGATGGGAGAGAGTGTGTCCAGTAAGAAATTA 360
Db GATTAAGATGAGTGGTGTGAGAGATGGGAGAGAGTGTGTCCAGTAAGAAATTA 68931
QY 361 GGCCTGCAACAGCTGAGAGGAGAGTGAAGAGAGAGAGGCGGAGATACAGATGA 420
Db GGCCTGCAACAGCTGAGAGGAGAGTGAAGAGAGAGAGGCGGAGATACAGATGA 68871
QY 421 GGGAGACAGCTGGAACAGAAAGTGAAGAGATTCAGATGTGAGAGAGAGGCTCA 480
Db GGGAGACAGCTGGAACAGAAAGTGAAGAGATTCAGATGTGAGAGAGAGGCTCA 68811
QY 481 CAGACCCCGCAATGATGTGTGACACAGAAATCTGGAAGAGAGAGATGATGAG 540
Db CAGACCCCGCAATGATGTGTGACACAGAAATCTGGAAGAGAGAGATGATGAG 68753
68810 CAG-GCCCCCGCAATGATGTGT-GACAAACAGAAATCTGGAAGAGAGATGATGAG 68753

QY 541 AGTGA CAAATGGGGTCTAAAGTTGAAGTTCAGAGCCAGGCAATGTGGCTACGCGCTGTA 600
AX331407/c
Db 68752 AGTGA CAAATGGGGTCTGAAGTTGAAGTTCAGAGCCAGGCAATGTAGCTACGCGCTGTA 68693
QY 601 ATCCCAACCTTTGAGGCTGAGAGTGGGGAATCACTTGAAGCCAGAGAGTTGAGACCGAG 660
Db 68692 ATCCCAATCTTTGAGGCTGAGAGTGGGGAATCACTTGAAGCCAGAGAGTTGAGACCGAG 68633
QY 661 CTTGGCCAACTGTGTGAAACCCCGTCTAC--AAAAAAAATACAAAAAATTTAGCCGGGGT 719
Db 68632 CTTGGCCAACTGTGTGAAACCCCGTCTACAAAAAAAATACAAAAAATTTAGCCGGGGT 68573
QY 720 TGGTATGAGACACTGTAGTCAAGCTACTTGGAGAGCTGAGGCAAGAGATTGCTTGA 779
Db 68572 TGGTATGAGACACTGTAGTCAAGCTACTTGGAGAGCTGAGGCAAGAGATTGCTTGA 68513
QY 780 CCGGGAGATGAGAGCTGAGAGTGTGAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 839
Db 68512 CCGGGAGATGAGAGCTGAGAGTGTGAGTCAAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 68453
QY 840 AGATGAAGCTCATCTCAAAAAAAGAGCTGATTTGAGTGAATATTATTAATCAAT 899
Db 68452 AGATGAAGCTCATCTC--AAAAAAGAGCTGATTTGAGTGAATATTATTAATCAAT 68395
QY 900 TCTCCCT 959
Db 68394 TCTCCCT 68335
QY 960 TCTGTAATTTCCATCTCTGTGTGTCTGTTCCTCATCTGCTTCTCATATTTGGGATCTCT 1019
Db 68334 TCTGTAATTTCCATCTCTGTGTGTCTGTTCCTCATCTGCTTCTCATATTTGGGATCTCT 68275
QY 1020 GGGTCTCATGTCTCTCTCTGCGCCACTTGGCCACTCTGCGCTCTCATGCGCCCGCT 1079
Db 68274 GGGTCTCATGTCTCTCTCTGCGCCACTTGGCCACTCTGCGCTCTCATGCGCCCGCT 68215
QY 1080 TTCTCTCTGAGAGGTATCTGTGGGGGCTGTGTCTGCAATGTGCTCTCTGAGGAGCT 1139
Db 68214 TTCTCTCTGAGAGGTATCTGTGGGGGCTGTGTGTGCAATGTGCTCTCTGAGGAGCT 68155
QY 1140 CGTGTCTGTGGGAGATTACCTTTGTGCGCGGCCAAAGACCGGGGTCTTACACGAACT 1199
Db 68154 CGTGTCTGTGGGAGATTACCTTTGTGCGCGGCCAAAGACCGGGGTCTTACACGAACT 68095
QY 1200 CTGCAAGTTCAACCAAGTGTATCAGAGAAACATCCAGGCACTCTGAGTCAATCCAGG 1259
Db 68094 CTGCAAGTTCAACCAAGTGTATCAGAGAAACATCCAGGCACTCTGAGTCAATCCAGG 68035
QY 1260 ACTGAGCACCGGCAATCCCACTGTCTGAGGAGACGCGCTGACACTCTCTTCAGAGCC 1319
Db 68034 ACTGAGCACCGGCAATCCCACTGTCTGAGGAGACGCGCGCTGACACTCTCTTCAGAGCC 67975
QY 1320 TCAATTCCTTCCAGAGATGTGAATGTTCATCTCTCAAGCCCTGACCCCATGTCTTC 1379
Db 67974 TCAATTCCTTCCAGAGATGTGAATGTTCATCTCTCAAGCCCTGACCCCATGTCTTC 67915
QY 1380 TGGAGCTCAGAGGTCTGTGCTCCCACTTGGGCTGACCGTGTCTCTCTAGTTGAACCTG 1439
Db 67914 TGGAGCTCAGAGGTCTGTGCTCCCACTTGGGCTGACCGTGTCTCTCTAGTTGAACCTG 67855
QY 1440 GAAACAATTTCAAAATGTCTGAGAGGCGGGGTTGTCTCAATCTCTCCGAGGCACTTTC 1499
Db 67854 GAAACAATTTCAAAATGTCTGAGAGGCGGGGTTGTCTCAATCTCTCCGAGGCACTTTC 67795
QY 1500 ATTCCTAAGCTCAGAGGCTCATCTCTCTCTGAGCTCTGACCAATTTAGTCCAGAAA 1559
Db 67794 ATTCCTAAGCTCAGAGGCTCATCTCTCTCTGAGCTCTGACCAATTTAGTCCAGAAA 67735
QY 1560 TAAACTGAGAG 1571
Db 67734 TAAACTGAGAG 67723

RESULT 9
AX331407/c
LOCUS AX331407 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 1916 from Patent WO0194629.
ACCESSION AX331407
VERSION AX331407.1 GI:18122041
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrihan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 1916 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
source location/Qualifiers
1..586
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 586;
Best Local Similarity 100.0%; Pred. No. 3,8e-111; Mismatches 0; Indels 0; Gaps 0;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1090 CAGGGTGAATTCCTGGGGGCTGTGTGTGCAATGAGCTCTCTGAGAGAGCTGTGCTCTGG 1149
Db 501 CAGGGTGAATTCCTGGGGGCTGTGTGTGCAATGAGCTCTCTGAGAGAGCTGTGCTCTGG 442
QY 1150 GGAAGTTACCTTTGGGCGCCGCAAGACCGGGTGTCTACAGAACTCTGCAAGTTC 1209
Db 441 GGAAGTTACCTTTGGGCGCCGCAAGACCGGGTGTCTACAGAACTCTGCAAGTTC 382
QY 1210 ACCAAGTGAATCAGAGAAACCATCCAGGCAACTCTGAGTCAATCCAGAGCTCAGACA 1269
Db 381 ACCAAGTGAATCAGAGAAACCATCCAGGCAACTCTGAGTCAATCCAGAGCTCAGACA 322
QY 1270 CCGGATCCCACTGTCTGAGAGGACAGCCCTGACACTCTTTCAGACCTTCATTCCTTC 1329
Db 321 CCGGATCCCACTGTCTGAGAGGACAGCCCTGACACTCTTTCAGACCTTCATTCCTTC 262
QY 1330 CCAAGATGTGAAGATGTTCATCTCTCAAGCCCTGACCCCATGTCTCTGAGACTCAGG 1389
Db 261 CCAAGATGTGAAGATGTTCATCTCTCAAGCCCTGACCCCATGTCTCTGAGACTCAGG 202
QY 1390 GTCTGCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTGAGAACTTC 1449
Db 201 GTCTGCTCCCACTTGGGCTGACCGTGTCTCTAGTTGAACCTGAGAACTTC 142
QY 1450 CAAAATCTGTCAAGGCGGGGTTGTGCTCAATCTCTCTGAGGCACTTTCATCTCAAGC 1509
Db 141 CAAAATCTGTCAAGGCGGGGTTGTGCTCAATCTCTCTGAGGCACTTTCATCTCAAGC 82
QY 1510 TCAAGGCCCATCTCTCTCTGAGCTGAGACCAATTTAGTCCAGAAATTAATCTGAGA 1569
Db 81 TCAAGGCCCATCTCTCTCTGAGCTGAGACCAATTTAGTCCAGAAATTAATCTGAGA 22
QY 1570 AG 1571
Db 21 AG 20
RESULT 10
AX331824/c
LOCUS AX331824 586 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 2333 from Patent WO0194629.
ACCESSION AX331824
VERSION AX331824.1 GI:18122458
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1
 Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G., Horrikan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature gene sets
JOURNAL Patent: WO 0194629-A 2333 13-DEC-2001;
 Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
 source
 1..586
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
ORIGIN
 Query Match 30.7%; Score 482; DB 6; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1090 CAGGGTGAATTCGGGGGGCTGTGATCTGCAATGGCTCCCTGAGGACTCGTCTCG 1149
 501 CAGGGTGAATTCGGGGGGCTGTGATCTGCAATGGCTCCCTGAGGACTCGTCTCG 442
 Qy 1150 GGAATTAACCTTGTGCTGCGGCGCAACAGACGGGTGTCTACAGAACTCTGCAAGTTC 1209
 Db 441 GGAATTAACCTTGTGCTGCGGCGCAACAGACGGGTGTCTACAGAACTCTGCAAGTTC 382
 Qy 1210 ACCAAGTGAATCCAGAAACATCCAGGCAACTCCGATGATCCAGGACTCAGACAC 1269
 Db 381 ACCAAGTGAATCCAGAAACATCCAGGCAACTCCGATGATCCAGGACTCAGACAC 322
 Qy 1270 CCGGATCCCACTCTGTCAGGAGACAGCCCTGACACTCTCTTACAGACCTCTTCTTC 1329
 Db 321 CCGGATCCCACTCTGTCAGGAGACAGCCCTGACACTCTCTTACAGACCTCTTCTTC 262
 Qy 1330 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 1389
 Db 261 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 202
 Qy 1390 GTCTGCTTCCCACTGAGGAGGCTGTGATCTCTAGTTGAACCTGGGAAACAATTC 1449
 Db 201 GTCTGCTTCCCACTGAGGAGGCTGTGATCTCTAGTTGAACCTGGGAAACAATTC 142
 Qy 1450 CAAACTGTCCAGGAGGAGGAGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1509
 Db 141 CAAACTGTCCAGGAGGAGGAGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGC 82
 Qy 1510 TCAGGAGGAGGAGGAGGAGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1569
 Db 81 TCAGGAGGAGGAGGAGGAGTGTGCTCAATCTCCCTGGGCACTTTCATCTCAAGC 22
 Qy 1570 AG 1571
 Db 21 AG 20

RESULT 11
LOCUS G41906 586 bp DNA linear STS 30-SEP-1998
DEFINITION SHGC-56840 Human Homo sapiens STS genomic, sequence tagged site.
ACCESSION G41906
VERSION G41906.1 GI:3668239
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 586)
AUTHORS Myers, R.M.
TITLE Human STS (1998)

JOURNAL Unpublished (1998)
COMMENT
 Contact: Richard M. Myers
 Stanford Human Genome Center (SHGC)
 Stanford University School of Medicine
 Department of Genetics, W-344, Stanford, CA 94305, USA.
 Tel: 4157259687
 Fax: 4157259689
 Email: myers@shgc.stanford.edu
 Primer A: AGAGACAGCTGACCCCAAT
 Primer B: GCCAATCTGATGATCTCC
 STS size: 188
 PCR Profile:
 Initial incubation: 95 degrees C for 10 minutes
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 30 seconds
 Polymerization: 72 degrees C for 23 seconds
 PCR Cycles: 30
 Thermal Cycler: Perkin Elmer 9700
 Template: 25 ng
 Primer: each 1 uM
 dNTPs: each 200 uM
 Amplifrag Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul
 Buffer: MgCl2: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3
 Prepared with primer pairs derived from W73140 -- Unigene.
FEATURES Location/Qualifiers
 source
 1..586
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="19"
 /clone_11b="Human"
 STS
 primer_bind
 167..186
 primer_bind
 complement(335..354)
ORIGIN
 Query Match 30.7%; Score 482; DB 11; Length 586;
 Best Local Similarity 100.0%; Pred. No. 3.8e-131;
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1090 CAGGGTGAATTCGGGGGGCTGTGATCTGCAATGGCTCCCTGAGGACTCGTCTCG 1149
 501 CAGGGTGAATTCGGGGGGCTGTGATCTGCAATGGCTCCCTGAGGACTCGTCTCG 442
 Qy 1150 GGAATTAACCTTGTGCTGCGGCGCAACAGACGGGTGTCTACAGAACTCTGCAAGTTC 1209
 Db 441 GGAATTAACCTTGTGCTGCGGCGCAACAGACGGGTGTCTACAGAACTCTGCAAGTTC 382
 Qy 1210 ACCAAGTGAATCCAGAAACATCCAGGCAACTCCGATGATCCAGGACTCAGACAC 1269
 Db 381 ACCAAGTGAATCCAGAAACATCCAGGCAACTCCGATGATCCAGGACTCAGACAC 322
 Qy 1270 CCGGATCCCACTCTGTCAGGAGACAGCCCTGACACTCTCTTACAGACCTCTTCTTC 1329
 Db 321 CCGGATCCCACTCTGTCAGGAGACAGCCCTGACACTCTCTTACAGACCTCTTCTTC 262
 Qy 1330 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 1389
 Db 261 CCAAGATGTTGAAATGTTTCTCTCCAGCCCTGACCCCAATGTTCTCTGAGCTCAG 202
 Qy 1390 GTCTGCTTCCCACTGAGGAGGCTGTGATCTCTAGTTGAACCTGGGAAACAATTC 1449
 Db 201 GTCTGCTTCCCACTGAGGAGGCTGTGATCTCTAGTTGAACCTGGGAAACAATTC 142

QY 1450 CAAAATGTCAGAGGGGGGTTGCGTTCATCTCCCTGAGGACATTTTCATCTCGAAG 1509
DB 141 CAAAATGTCAGAGGGGGGTTGCGTTCATCTCCCTGAGGACATTTTCATCTCGAAG 82
QY 1510 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569
DB 81 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 22
QY 1570 AG 1571
DB 21 AG 20

RESULT 12
AX429955/c 735 bp DNA linear PAT 21-JUN-2002
LOCUS AX429955
DEFINITION Sequence 94 from Patent WO0198339.
ACCESSION AX429955
VERSION AX429955.1 GI:21541119
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.
TITLE Compositions and methods for the therapy and diagnosis of breast cancer
JOURNAL Patent: WO 0198339-A 94 27-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source 1..735
location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 735;
Best Local Similarity 100.0%; Pred. No. 3.9e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCTGTCTCG 1149
DB 500 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCTGTCTCG 441
QY 1150 GGAGATTACCTTTGTGCTCGGCGCCCAACAGCGGGTGTCTACGCAACTCTGCAAGTTC 1209
DB 440 GGAGATTACCTTTGTGCTCGGCGCCCAACAGCGGGTGTCTACGCAACTCTGCAAGTTC 381
QY 1210 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 1269
DB 380 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 321
QY 1270 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTACAGACCTCATTCCTTC 1329
DB 320 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTACAGACCTCATTCCTTC 261
QY 1330 CCAAGATTTGAAGATTCTCTCCAGCCCTGACCCCAAGTCTCTGAGACTGAG 1389
DB 260 CCAAGATTTGAAGATTCTCTCCAGCCCTGACCCCAAGTCTCTGAGACTGAG 201
QY 1390 GTCTGTTCCCACTGGGGTGAACCGTGTCTCTAGTTGAACCGTGGGAACTTTC 1449
DB 200 GTCTGTTCCCACTGGGGTGAACCGTGTCTCTAGTTGAACCGTGGGAACTTTC 141
QY 1450 CAAAATGTCAGAGGGGGGTTGCGTTCATCTCCCTGAGGACATTTTCATCTCGAAG 1509
DB 140 CAAAATGTCAGAGGGGGGTTGCGTTCATCTCCCTGAGGACATTTTCATCTCGAAG 81
QY 1510 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 1569
DB 80 TCAGGGCCCATCCCTTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 21

QY 1570 AG 1571
DB 20 AG 19

RESULT 13
BD139877/c 738 bp DNA linear PAT 18-SEP-2002
LOCUS BD139877
DEFINITION Compounds for immunotherapy and diagnosis of breast cancer and methods for their use.
ACCESSION BD139877
VERSION BD139877.1 GI:23234822
KEYWORDS JP 2002507387-A/94.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Reed,S.G. and Xu,J.
TITLE Compounds for immunotherapy and diagnosis of breast cancer and methods for their use
JOURNAL Patent: JP 2002507387-A 94 12-MAR-2002;
CORIXA CORP
COMMENT OS Homo sapiens (human)
PN JP 2002507387-A/94
PD 12-MAR-2002
PF 22-DEC-1998 JP 2000526543
PR 24-DEC-1997 US 08/998253, 24-DEC-1997 US 08/998255 PR
17-JUL-1998 US 09/118627, 17-JUL-1998 US 09/118554 PI STEVEN
G REED, JIANGCHUN XU
PC C12N15/09,A61K38/00,A61K39/00,A61K39/39,A61K39/395,A61P35/00,
PC C07K14/47,
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/PC
53,
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00
CC Compounds for immunotherapy and diagnosis
of breast cancer and
CC their use methods for
CC key Location/Qualifiers
FT source 1..738
FT /organism="Homo sapiens"
location/Qualifiers
1..738
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 738;
Best Local Similarity 100.0%; Pred. No. 3.9e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCTGTCTCG 1149
DB 503 CAGGGTATCTGGGGGGCTGTGCTGCAATGCTCCCTGAGGACTCTGTCTCG 444
QY 1150 GGAGATTACCTTTGTGCTCGGCGCCCAACAGCGGGTGTCTACGCAACTCTGCAAGTTC 1209
DB 443 GGAGATTACCTTTGTGCTCGGCGCCCAACAGCGGGTGTCTACGCAACTCTGCAAGTTC 384
QY 1210 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 1269
DB 383 ACCAAGTATCCAGGAAACATCCAGGCACTCCTGAGTATCCAGGACTCGACACA 324
QY 1270 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTACAGACCTCATTCCTTC 1329
DB 323 CCGGATCCCACTGCTGCAAGGACAGCCCTGACACTCTCTTACAGACCTCATTCCTTC 264
QY 1330 CCAAGATTTGAAGATTCTCTCCAGCCCTGACCCCAAGTCTCTGAGACTGAG 1389
DB 260 CCAAGATTTGAAGATTCTCTCCAGCCCTGACCCCAAGTCTCTGAGACTGAG 201

Db 263 CCAGAGATGTGAAGATGTTTCATCTCTCCAGCCCTGACCCCAATGTCCTGAGCTCAG 204
Qy 1390 GTCTGCTTCCCAATGTTGGGCTGACCGGTGTCTCTAGTTGAACCTGGGAACAATTTTC 1449
Db 203 GTCTGCTTCCCAATGTTGGGCTGACCGGTGTCTCTAGTTGAACCTGGGAACAATTTTC 144
Qy 1450 CAAAATGTCAGAGGGGGGGGGTTCGTCATCTCCCTGGGGAATTTATCTCCTCAAGC 1509
Db 143 CAAAATGTCAGAGGGGGGGGGTTCGTCATCTCCCTGGGGAATTTATCTCCTCAAGC 84
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 83 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 24
Qy 1570 AG 1571
Db 23 AG 22

RESULT 14
LOCUS CQ72600 1143 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 6534 from Patent WO02068579.
ACCESSION CQ72600
VERSION CQ72600.1 GI:42281457
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 6534 06-SEP-2002;
JOURNAL PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1. .1143
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 30.7%; Score 482; DB 6; Length 1143;
Best Local Similarity 100.0%; Pred. No. 4e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGGATCTTGGGGGGGCTGTGTGCAATGGCTCCCTGACAGGGACCTGTCCTGG 1149
Db 657 CAGGGGATCTTGGGGGGGCTGTGTGCAATGGCTCCCTGACAGGGACCTGTCCTGG 716
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 1209
Db 717 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 776
Qy 1210 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTGAGTCAATCCAGAGACTCAGACA 1269
Db 777 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTGAGTCAATCCAGAGACTCAGACA 836
Qy 1270 CCGGCAATCCCACTGCTGTGAGAGGACGCTGACATCTCTTCAAGCCTCAATTCCTTC 1329
Db 837 CCGGCAATCCCACTGCTGTGAGAGGACGCTGACATCTCTTCAAGCCTCAATTCCTTC 896
Qy 1330 CCAGAGATGTGAAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCCTGAGACTCAG 1389
Db 897 CCAGAGATGTGAAGATGTTTCATCTCTCAGCCCTGACCCCAATGTCCTGAGACTCAG 956
Qy 1390 GTCTGCTTCCCAATGTTGGGCTGACCGGTGTCTCTAGTTGAACCTGGGAACAATTTTC 1449
Db 957 GTCTGCTTCCCAATGTTGGGCTGACCGGTGTCTCTAGTTGAACCTGGGAACAATTTTC 1016
Qy 1450 CAAAATGTCAGAGGGGGGGGGTTCGTCATCTCCCTGGGGAATTTATCTCCTCAAGC 1509

Db 1017 CAAAATGTCAGAGGGGGGGGGTTCGTCATCTCCCTGGGGAATTTATCTCCTCAAGC 1076
Qy 1510 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1569
Db 1077 TCAGGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCAGAAATAAAGTGA 1136
Qy 1570 AG 1571
Db 1137 AG 1138

RESULT 15
LOCUS AY461805 1301 bp mRNA linear PRI 03-DEC-2003
DEFINITION Homo sapiens kallikrein 5 isoform 3 preproprotein (KLK5) mRNA,
complete cds; alternatively spliced.
ACCESSION AY461805
VERSION AY461805.1 GI:38564772
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
Michael, I.P., Kurlender, L., Du, D.C. and Diamond, E.P.
Cloning of new splice variants of the human kallikrein gene 5
2 (bases 1 to 1301)
REFERENCE 2
Michael, I.P., Kurlender, L., Du, D.C. and Diamond, E.P.
Direct Submission
Submitted (10-NOV-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
JOURNAL
FEATURES
source Location/Qualifiers
1. .1301
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. .1301
/gene="KLK5"
12..626
/gene="KLK5"
/note="alternatively spliced"
/codon_start=1
/product="kallikrein 5 isoform 3 preproprotein"
/protein_id="AAR23814.1"
/db_xref="GI:38564773"
/translation="MATARPMMWVLCALITLILGTVRHVLANDVSCDHPSTVPS
GSNDIGAGAGDARDSDSSRIINGSDCDMTTPQALILRPNDLYCGAVLVHFW
LTLAAHCRKVRVRLGSHLSPVESGQMGQVYSIHPGVSHGHSDMLIKLN
RRIRPKDVRPINTVSHSPSAGTKCLVSGMTTKSPQSGPSS"

ORIGIN
Query Match 30.7%; Score 482; DB 9; Length 1301;
Best Local Similarity 100.0%; Pred. No. 4.1e-131;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1090 CAGGGGATCTTGGGGGGGCTGTGTGCAATGGCTCCCTGACAGGGACCTGTCCTGG 1149
Db 820 CAGGGGATCTTGGGGGGGCTGTGTGCAATGGCTCCCTGACAGGGACCTGTCCTGG 879
Qy 1150 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 1209
Db 880 GGAGATTACCTTGTGCGCCGCCCAACAGACCGGGGTCTTACAGAACTCTGCAATTC 939
Qy 1210 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTGAGTCAATCCAGAGACTCAGACA 1269
Db 940 ACCAAGTGATCCAGGAACCAATCCAGGCCCACTCTGAGTCAATCCAGAGACTCAGACA 999
Qy 1270 CCGGCAATCCCACTGCTGTGAGAGGACGCTGACATCTCTTCAAGCCTCAATTCCTTC 1329

```
Db      1000  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY      1330  CCAGAGATGTTGAAATGTTCAATCTCTCAGAGCCCTGACACTCCTTTCAGACCTCATTCCTTC
Db      1060  |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY      1390  GTCTGCTTCCCGCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGGAACAATTTC
Db      1120  GTCTGCTTCCCGCACAATTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGGAACAATTTC
QY      1450  CAAAACTGTTCAGAGGCGGGGCTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC
Db      1180  CAAAACTGTTCAGAGGCGGGGCTTGCGTCTCAATCTCCCTGGGGCACTTTCATCTCAAGC
QY      1510  TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTAGTCCAGAAATAAATGAGA
Db      1240  TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTAGTCCAGAAATAAATGAGA
QY      1570  AG 1571
Db      1300  AG 1301
```

Search completed: February 26, 2005, 06:24:09
Job time : 6702.88 secs

1	5000	100.0	11570	6	C07882319	Sequence
2	5000	100.0	11570	6	C0874885	Sequence
3	5000	100.0	11570	6	C0874961	Sequence
4	5000	100.0	11570	6	AF135028	Homo sapi
5	4989	99.8	107487	9	AF135028	Homo sapi
6	4989	99.8	230000	9	AC011483	Homo sapi
7	4243.4	84.9	217346	9	AF243527	Homo sapi
8	4106.4	82.1	200752	2	AC027602	Homo sapi
9	1035.2	20.7	176647	2	AC130782	Pat. tlog
10	451.6	9.2	32789	9	AC130188	Homo sapi
11	457.2	9.1	190603	9	AC137591	Homo sapi
12	456.6	9.1	141923	2	AC149102	Homo sapi
13	455	9.1	138849	9	AC105997	Homo sapi
14	455	9.1	171637	9	HS95C30	Human DNA
15	455	9.1	127137	9	AC114876	Homo sapi
16	452.4	9.0	119147	9	AC098481	Homo sapi
17	451.6	9.0	210508	2	AC135612	Homo sapi
18	450.6	9.0	100893	3	AC079169	Homo sapi
19	450	9.0	100208	9	AC135468	Homo sapi

45	437.6	8.8	129711	9	AC110768	Homo sapi
44	438	8.8	166666	9	AC103895	Homo sapi
43	438	8.8	95930	9	AC106033	Homo sapi
42	439	8.8	90067	9	AL139234	Human DNA
41	439	8.8	135640	9	AC136718	Homo sapi
40	439	8.8	45866	9	HSBA9P11	Human DNA
39	439.4	8.8	50054	9	AC093691	Homo sapi
38	439.6	8.8	173058	9	AC104564	Homo sapi
37	440.4	8.8	81278	9	AC1091062	Homo sapi
36	441.4	8.8	158003	9	AC022795	Homo sapi
35	442	8.8	263876	9	AC012933	Homo sapi
34	442	8.8	109907	9	AC092842	Homo sapi
33	442.4	8.8	198888	9	AC087728	Homo sapi
32	442.4	8.8	196662	9	AP001107	Homo sapi
31	442.4	8.8	193159	9	AC006546	Homo sapi
30	442.4	8.8	101269	9	AC005300	Homo sapi
29	443.4	8.9	106277	9	AC068792	Homo sapi
28	444	8.9	246282	9	HSU91328	Human hereid
27	444	8.9	246240	6	AR036574	Sequence
26	444	8.9	266240	6	AR036574	Sequence
25	444	8.9	246240	6	AR036572	Sequence
24	447.4	8.9	99876	9	HSB199108	Human DNA
23	449.4	9.0	183660	9	AC091180	Homo sapi
22	449.6	9.0	147122	9	AC090109	Homo sapi
21	450	9.0	119555	9	EX004627	Human DNA
20	450.8	9.0	182569	9	AC092500	Homo sapi

ALIGNMENTS

	RESULT 1								
LOCUS	CQ788219	11570 bp	DNA	linear	PAT 24-MAR-2006				
DEFINITION	Sequence 2 from Patent WO2004021008.								
ACCESSION	CQ788219								
VERSION	CQ788219.1	GI:45723068							
KEYWORDS									
SOURCE	Homo sapiens (human)								
ORGANISM	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.								
REFERENCE	1 Diamond J.E.P. Methods for detecting breast and ovarian cancer Patent: WO 2004021008-A 2 11-MAR-2004;								
AUTHORS	MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)								
TITLE	Location/Qualifiers								
JOURNAL	1 . 11570 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"								
FEATURES									
SOURCE									
ORIGIN									
Query Match	100.0%; Score 5000; DB 6; Length 11570;								
Best Local Similarity	100.0%; Fred. No. 0;								
Matches 5000; Conservative	0; Mismatches 0; Indels 0; Gaps 0;								
Dy	1 GGCGCCAGAGTGAAGGCAAGAAGAAGATTGAGACTCCTCTGCMAAAGTGCGTTGAGTC	60							
Dd	1 GGCGCCAGAGTGAAGGCAAGAAGAAGATTGAGACTCCTCTGCMAAAGTGCGTTGAGTC	60							
Dy	61 TCCCCTTGGCTTAATAATGCCAGGAGAGGGAGGCAAAAGAACAAGAGAGAGGGGTGGGG	120							
Dd	61 TCCCCTTGGCTTAATAATGCCAGGAGAGGGAGGCAAAAGAACAAGAGAGAGGGGTGGGG	120							
Dy	121 AAGAAAAGAGAGAGAGAGAGAGACAGATTACACAACTACAGAAACACAGAGAGAAC	180							
Dd	121 AAGAAAAGAGAGAGAGAGAGAGACAGATTACACAACTACAGAAACACAGAGAGAAC	180							
Dy	181 ACAGAGAGCCTGGGACACAGGGACACACAGAGTCAAGAGAAAAAGAGAGATTAGAGAA	240							
Dd	181 ACAGAGAGCCTGGGACACAGGGACACACAGAGTCAAGAGAAAAAGAGAGATTAGAGAA	240							

OY	241	ICACAAATGAGACACACAGGGGTAAATAAAGAGATTAAACAGTCCCAATACACG	300
Db	241	ACACAAATGAGACACACAGGGGTAAATAAAGAGATTAAACAGTCCCAATACACG	300
OY	301	AAAGGGGACAGACACAGTTTTACAGGGGAGTGTCTATCATCATCTTCTTTTTTTTTT	360
Db	301	AAAGGGGACAGACACAGTTTTACAGGGGAGTGTCTATCATCATCTTCTTTTTTTTTT	360
OY	361	TTTTTTTTTTTTTTTTTGACGAGTCTGGCTCTGTCCGCAAGGCTGAGTGCATGGCG	420
Db	361	TTTTTTTTTTTTTTTTTGACGAGTCTGGCTCTGTCCGCAAGGCTGAGTGCATGGCG	420
OY	421	GATCTCGGCTCACTCAGACTCCGCGCTCCGGGTTCAAGCATTTCTCTGCTCAGCTC	480
Db	421	GATCTCGGCTCACTCAGACTCCGCGCTCCGGGTTCAAGCATTTCTCTGCTCAGCTC	480
OY	481	CCAAAGTACTGGAGCTACAGGCGCCGCACTACCCCGGCTAAATTTTTGTATTTTAA	540
Db	481	CCAAAGTACTGGAGCTACAGGCGCCGCACTACCCCGGCTAAATTTTTGTATTTTAA	540
OY	541	GTAAGAGAGGGGGTTTCACTGTTTACCGGGATGGGCTGATCTCTACCTCGATCC	600
Db	541	GTAAGAGAGGGGGTTTCACTGTTTACCGGGATGGGCTGATCTCTACCTCGATCC	600
OY	601	GCCGCGCTCGGCTCCCAAAGTGTCTGGATTACAGGCGTGAAGCACCAGCGCCGATG	660
Db	601	GCCGCGCTCGGCTCCCAAAGTGTCTGGATTACAGGCGTGAAGCACCAGCGCCGATG	660
OY	661	ATCATCTTCTTGACATATGCTGATGTGACAAATACCTAAAGCATACCTCACTTTTA	720
Db	661	ATCATCTTCTTGACATATGCTGATGTGACAAATACCTAAAGCATACCTCACTTTTA	720
OY	721	AATATGCAAGTTTGGGCGACGACACGCTGCTCATGCTCTAATTCAGACTTTGGAGGC	780
Db	721	AATATGCAAGTTTGGGCGACGACACGCTGCTCATGCTCTAATTCAGACTTTGGAGGC	780
OY	781	AGAGTGGGTGATCATCTTGAGGCGCAGAGTTTGAGACCAAGCTTGGCCAAATGTTGAA	840
Db	781	AGAGTGGGTGATCATCTTGAGGCGCAGAGTTTGAGACCAAGCTTGGCCAAATGTTGAA	840
OY	841	CTGTGCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	900
Db	841	CTGTGCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	900
OY	901	ACCTGTAATCCCAAGTATGCTGAGGCGTGAAGCAGAGATCATCTTGAACCTGTGAGCG	960
Db	901	ACCTGTAATCCCAAGTATGCTGAGGCGTGAAGCAGAGATCATCTTGAACCTGTGAGCG	960
OY	961	GAGGTGAGTGGGCGCGAGATCACTACCCGCTTCAGCTGGGCGACAAGCAAGCT	1020
Db	961	GAGGTGAGTGGGCGCGAGATCACTACCCGCTTCAGCTGGGCGACAAGCAAGCT	1020
OY	1021	CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT	1080
Db	1021	CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT	1080
OY	1081	AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1140
Db	1081	AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1140
OY	1141	AGAACTCTAAGGTATATTGACAAATTCATTCAGAACTTTTAAAAAAGAAAGATACAGA	1200
Db	1141	AGAACTCTAAGGTATATTGACAAATTCATTCAGAACTTTTAAAAAAGAAAGATACAGA	1200
OY	1201	GGCATAGAAAGACAGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAAACAA	1260
Db	1201	GGCATAGAAAGACAGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAAACAA	1260
OY	1261	ACAAAGGCTCTTAAGCACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Db	1261	ACAAAGGCTCTTAAGCACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320

QY	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGCCGAGGGGATPAGAAAG	1380
Db	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGCGAGGGGATPAGAAAG	1380
QY	1381	GAGAGAGGGGTGAGAGAGACAGAGATTTTGGAGAGACTCAGAAAGATAGCCGAGGGA	1440
Db	1381	GAGAGAGGGGTGAGAGAGACAGAGATTTTGGAGAGACTCAGAAAGATAGCCGAGGGA	1440
QY	1441	GAAACAAGAGAGATGGAAAGAAAGCTCTGAGAAAAAAACAAGACAAAGATGAAAGAGG	1500
Db	1441	GAAACAAGAGAGATGAAAGAAAGCTCTGAGAAAAAAACAAGACAAAGATGAAAGAGG	1500
QY	1501	AGTATCGAGGGTGAACAGACAGTGGTGGAAATGACAAATTCGAGAAAGAAACCAAGCAA	1560
Db	1501	AGTATCGAGGGTGAACAGACAGTGGTGGAAATGACAAATTCGAGAAAGAAACCAAGCAA	1560
QY	1561	TCACAGGCGCCAGAGATATGACCCAGAGATTGGTGAAGAACCAAGATCCTTAAGCGTGGGG	1620
Db	1561	TCACAGGCGCCAGAGATATGACCCAGAGATTGGTGAAGAACCAAGATCCTTAAGCGTGGGG	1620
QY	1621	AGGCAGGAAAGGGGCTGGGCTTTCGGAGACCCCTCCCAATTCCTCGGGCCAGGGAG	1680
Db	1621	AGGCAGGAAAGGGGCTGGGCTTTCGGAGACCCCTCCCAATTCCTCGGGCCAGGGAG	1680
QY	1681	GTAAGGAGTGAATTCCTCGGACTGGGTGGGGGGTCTCTTGGGGGTGAGAGATAGGGGAGCA	1740
Db	1681	GTAAGGAGTGAATTCCTCGGACTGGGTGGGGGGTCTCTTGGGGGTGAGAGATAGGGGAGCA	1740
QY	1741	GGAAGAGCTAATTGCTAAGGCCCGATAGGACCTCATTTGCCGGGAAATGTGCCCAAGAG	1800
Db	1741	GGAAGAGCTAATTGCTAAGGCCCGATAGGACCTCATTTGCCGGGAAATGTGCCCAAGAG	1800
QY	1801	CAGTGGGTGTATTAATCTCAGGCCCGGTGCCAGAGCCCAAGAGAGCGAGTGGCCAGGA	1860
Db	1801	CAGTGGGTGTATTAATCTCAGGCCCGGTGCCAGAGCCCAAGAGAGAGCAGTGGCCAGGA	1860
QY	1861	AGGCACAGGCCCTGAGAAAGTCTGCGGCTGAGCTGGAGGCAATTCGCCCAACCTTACTCTGG	1920
Db	1861	AGGCACAGGCCCTGAGAAAGTCTGCGGCTGAGCTGGAGGCAATTCGCCCAACCTTACTCTGG	1920
QY	1921	GCGACAGGGCAAGTGAAGACTGTGTGAGGGTGGCTCAGCAGGCCAGGGAAAGAGAGTGTCT	1980
Db	1921	GCGACAGGGCAAGTGAAGACTGTGTGAGGGTGGCTCAGCAGGCCAGGGAAAGAGAGTGTCT	1980
QY	1981	GTGCGTCCCTGACCCACATCTTTCTCTGTCCCCCTCTTGCCTGTCTGAGGCTGTAGA	2040
Db	1981	GTGCGTCCCTGACCCACATCTTTCTCTGTCCCCCTCTTGCCTGTCTGAGGCTGTAGA	2040
QY	2041	CTCCATATCTTCTGAATTTCTATAGTGTGCGGTCTCAGCCGACAGTGTGTGGCCGCTC	2100
Db	2041	CTCCATATCTTCTGAATTTCTATAGTGTGCGGTCTCAGCCGACAGTGTGTGGCCGCTC	2100
QY	2101	CTTGTGGTTCCCTCTACCTGGGGAAATTAAGGTAGGGGAGGGAGGGGAAAGTGGATTAAAG	2160
Db	2101	CTTGTGTTCCTCTCTACTGTGGGAAATTAAGGTAGGGGAGGGAGGGGAAAGTGGATTAAAG	2160
QY	2161	GCTCCCGGATGCGCTGGGCTCCCAACCTCTGAATTCGCCATCAGGTGACAGCGCC	2220
Db	2161	GCTCCCGGATGCGCTGGGCTCCCAACCTCTGAATTCGCCATCAGGTGACAGCGCC	2220
QY	2221	ATGGCTACAGCAAGACCCCCCTGAGATGTGGGTCTCTGTGCTCTGATTCACAGCTTGCTT	2280
Db	2221	ATGGCTACAGCAAGACCCCCCTGAGATGTGGGTCTCTGTGCTCTGATTCACAGCTTGCTT	2280
QY	2281	CTGGGGGTGACAGGTAAACAGAACTCTGGGGTGGAGGGGTGTGGGAATTTGGAGAGACTGT	2340
Db	2281	CTGGGGGTGACAGGTAAACAGAACTCTGGGGTGGAGGGGTGTGGGAATTTGGAGAGACTGT	2340
QY	2341	CTCTCGGCACTAAGAGCGCTGTCCCTTGGGAACTGTGTGAGCCTTGGGATGACTCTCGG	2400
Db	2341	CTCTCGGCACTAAGAGCGCTGTCCCTTGGGAACTGTGTGAGCCTTGGGATGACTCTCGG	2400
QY	2401	GACCGGATGATGTGATCTCTGTCTGATCTTGTGGTGTGCGATCGTATGTGGCCCTGT	2460

Db	2401	GACCGGGGAATGATAGTCTCTGTGTACTTGTTGGTTGTGGATCGATATGGCCCTGT	2466
Oy	2461	GACTGCCACGGGTGTGTTCGGGGAGGGGGAATGCCCTTTTCCAAATACAGGTGACTGTGC	2520
Db	2461	GACTGCCACGGGTGTGTTCGGGGAGGGGGAATGCCCTTTTCCAAATACAGGTGACTGTGC	2520
Oy	2521	CAGGTGGCACTGACCCTTTGAGAAGCTGTGTGTGTGGTTTTGTGATTGTGTGCATTTAAG	2580
Db	2521	CAGGTGGCACTGACCCTTTGAGAAGCTGTGTGTGTGGTTTTGTGATTGTGTGCATTTAAG	2580
Oy	2581	AATGTGTGTGGCTCCACAAGCTGTGTGGGTGAATGCAATGCACATGGGGGGTTCACCTGT	2640
Db	2581	AATGTGTGTGGCTCCACAAGCTGTGTGGGTGAATGCAATGCACATGGGGGGTTCACCTGT	2640
Oy	2641	GTGTGTGGCTGTGTGTGTGACTTTGGCAATTGTAATATGACTCAGGATATCTCAGTTCCGTG	2700
Db	2641	GTGTGTGGCTGTGTGTGTGACTTTGGCAATTGTAATATGACTCAGGATATCTCAGTTCCGTG	2700
Oy	2701	TCCCTGAGAGTCCCGGGAATTCGCTGCACAAAAGTGTGATCACCATGGAAAAGCTGTGACT	2760
Db	2701	TCCCTGAGAGTCCCGGGAATTCGCTGCACAAAAGTGTGATCACCATGGAAAAGCTGTGACT	2760
Oy	2761	GTGTGTGTCTTTCAGGCGAATTATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2820
Db	2761	GTGTGTGTCTTTCAGGCGAATTATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2820
Oy	2821	TTTGTGACCGGTGTGACTACCTGAGAGCTCGTGTGTAGGGGTGACTGTATGTGACTGTGTGTG	2880
Db	2821	TTTGTGACCGGTGTGACTACCTGAGAGCTCGTGTGTAGGGGTGACTGTATGTGACTGTGTGTG	2880
Oy	2881	TCTGTGTAGGCGCGTGTAAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2940
Db	2881	TCTGTGTAGGCGCGTGTAAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2940
Oy	2941	TTTGTGTCTGT	3000
Db	2941	TTTGTGTCTGT	3000
Oy	3001	CAGGTGACTGACTTTCAGT	3060
Db	3001	CAGGTGACTGACTTTCAGT	3060
Oy	3061	TGTGTGACACAGGCAATCTGT	3120
Db	3061	TGTGTGACACAGGCAATCTGT	3120
Oy	3121	GCTAGAGCTGT	3180
Db	3121	GCTAGAGCTGT	3180
Oy	3181	GAGGCAATGGGCGGTGTCTGTGAGAACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3240
Db	3181	GAGGCAATGGGCGGTGTCTGTGAGAACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3240
Oy	3241	GCGTGTGTCTTGTGGGT	3300
Db	3241	GCGTGTGTCTTGTGGGT	3300
Oy	3301	GAAACGAGCCCGGCGCGGT	3360
Db	3301	GAAACGAGCCCGGCGCGGT	3360
Oy	3361	GGCGGATACCTTAGGT	3420
Db	3361	GGCGGATACCTTAGGT	3420
Oy	3421	CTACTATAAAATACAAAAATTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3480
Db	3421	CTACTATAAAATACAAAAATTAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3480
Oy	3481	GGAGACTGTGGGACAAAAATCGCTTGAACCCGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3540
Db	3481	GGAGACTGTGGGACAAAAATCGCTTGAACCCGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3540

D	b		3481	GGGAGACTGGGCGCAGAAAAATTCGTTGAACCCGGGAGAGTGGAGGTTGCGGTGACCCAGA	3540
O	y		3541	TCGGCCATTGCACTCCAGCCTGGGCAACAAGACGAACTCGGTCTTGAAAGAAAAA	3600
D	b		3541	TCGGCCATTGCACTCCAGCCTGGGCAACAAGACGAACTCGGTCTTGAAAGAAAAA	3600
O	y		3601	GAIAAAAAAAAAAGGTTAAGAACCACTGAATGGGACACGGGAGACTGAATGATGGAATTGGGCG	3666
D	b		3601	GAIAAAAAAAAAAGGTTAAGAACCACTGAATGGGACACGGGAGACTGAATGATGGAATTGGGCG	3666
O	y		3661	ATGCATGATGACTGTAGAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAAGGCAT	3720
D	b		3661	ATGCATGATGACTGTAGAGTCTGTGTGAGAGGAGAGATTGACAGGATTGAGAAAGGCAT	3720
O	y		3721	GTTTTCAATCTGAGAAATTCAGAAACTTAGCGCTGCTCTCCCTGCATGATGGCCCCCTTAAG	3780
D	b		3721	GTTTTCAATCTGAGAAATTCAGAAACTTAGCGCTGCTCTCCCTGCATGATGGCCCCCTTAAG	3780
O	y		3781	CTGAGCCCTTCTTCTGCTGCTGCTGCTCTGCTTCCGAACTTAAGTCCGACATGAGCTTGACC	3840
D	b		3781	CTGAGCCCTTCTTCTGCTGCTGCTGCTCTGCTTCCGAACTTAAGTCCGACATGAGCTTGACC	3840
O	y		3841	CCACTCTCTTCCCTCAACACAGCCGCTAAGCANAACCTTAGTGGACCCCGCTAAGGCGCA	3900
D	b		3841	CCACTCTCTTCCCTCAACACAGCCGCTAAGCANAACCTTAGTGGACCCCGCTAAGGCGCA	3900
O	y		3901	CACCCCTTTGGGCGCAGGCTCCACCCCTAATCTGTGGGTACTCTTCTAGAAACCCCTTCAA	3966
D	b		3901	CACCCCTTTGGGCGCAGGCTCCACCCCTAATCTGTGGGTACTCTTCTAGAAACCCCTTCAA	3966
O	y		3961	AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTGCCAGCTGGAG	4020
D	b		3961	AGTCAGAGCTTTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTTGCTCTCTGCCAGCTGGAG	4020
O	y		4021	TGCGATGGCGGATCTCGGCTCACTGGAACCTCGCTCCAGGTTAAAGATTTCTCGT	4086
D	b		4021	TGCGATGGCGGATCTCGGCTCACTGGAACCTCGCTCCAGGTTAAAGATTTCTCGT	4086
O	y		4081	GCGTCCACCTCTGATGATGCTGGATTAACAGGTGCGGCGCACACAGCGCTGGCTAATTTTT	4140
D	b		4081	GCGTCCACCTCTGATGATGCTGGATTAACAGGTGCGGCGCACACAGCGCTGGCTAATTTTT	4140
O	y		4141	GTCCTTTAGTAGAGACAGGGGTTTCACTTGTTGGCAGGCTGGTCTCAAACTCCAAAC	4200
D	b		4141	GTCCTTTAGTAGAGACAGGGGTTTCACTTGTTGGCAGGCTGGTCTCAAACTCCAAAC	4200
O	y		4201	TCAAGTGAATCCGCCCACTCGGCTCCCAAGATGCTGGGTAAAGCGTAGGACACCGC	4266
D	b		4201	TCAAGTGAATCCGCCCACTCGGCTCCCAAGATGCTGGGTAAAGCGTAGGACACCGC	4266
O	y		4261	CCCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAACATGTAAACCTTGACCTTGACC	4320
D	b		4261	CCCCAGGCCAAAGTCAGAGCTCTTTATAGGAGACTTAACATGTAAACCTTGACCTTGACC	4320
O	y		4321	CTAACTAAGTCAATTCCAAACCCCTTCTGCTGCACACCCCTGACCTCACTGAGGCGC	4380
D	b		4321	CTAACTAAGTCAATTCCAAACCCCTTCTGCTGCACACCCCTGACCTCACTGAGGCGC	4380
O	y		4381	TGACCCCACTTTGAGAACCAAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCACTCCGAG	4440
D	b		4381	TGACCCCACTTTGAGAACCAAGTTCATCCCTAAAGCCCTGGTCTCCCTCCCACTCCGAG	4440
O	y		4441	GCTCAGGCCCCCAGAGCTTTGGCACTAACCCCTGAGCTTGTCCAGAGATTCGTACCCAAAT	4500
D	b		4441	GCTCAGGCCCCCAGAGCTTTGGCACTAACCCCTGAGCTTGTCCAGAGATTCGTACCCAAAT	4500
O	y		4501	TTTTACCCCTCAATGTAGTTCTTAGCCAAATTCAGGAATCTGTAGGTTCCAGTTAGATGCA	4560
D	b		4501	TTTTACCCCTCAATGTAGTTCTTAGCCAAATTCAGGAATCTGTAGGTTCCAGTTAGATGCA	4560
O	y		4561	GTAACCCCTACTGAGCTGGGCTCTGTCTTGTAGACTTGAAGCTGGGCTGTAGAGATGCCA	4620
D	b		4561	GTAACCCCTACTGAGCTGGGCTCTGTCTTGTAGACTTGAAGCTGGGCTGTAGAGATGCCA	4620

QY 4621 CTCTTATTCAGAGCCCTGCCCCCTCAGCATGTCAGACACCCCTTAGCT 4680
DB 4621 CTCTTATTCAGAGCCCTGCCCCCTCAGCATGTCAGACACCCCTTAGCT 4680
QY 4681 GGTCTGCTCTTGAAGTCTGAAACCAACCCCAAGCCCGCTTGAAGCCCGC 4740
DB 4681 GGTCTGCTCTTGAAGTCTGAAACCAACCCCAAGCCCGCTTGAAGCCCGC 4740
QY 4741 CAACCATTTTCCGTTCCAGAGCATGTTCTGSCCAACATGATGTTTCTGACAC 4800
DB 4741 CAACCATTTTCCGTTCCAGAGCATGTTCTGSCCAACATGATGTTTCTGACAC 4800
QY 4801 CCTTAAACCGTCCCTCTGAGAGCAACAGACCTGAGAGCTGAGGCGGAGAGAG 4860
DB 4801 CCTTAAACCGTCCCTCTGAGAGCAACAGACCTGAGAGCTGAGGCGGAGAGAG 4860
QY 4861 CCCGCTGGATGACAGACAGACCGCATCATCATGATCCGATGATGACACCC 4920
DB 4861 CCCGCTGGATGACAGACAGACCGCATCATCATGATCCGATGATGACACCC 4920
QY 4921 AGCCGTGGAGCGGCTGTTGCTAAGGCCCAACAGCTTACGCGGCGGCTTGG 4980
DB 4921 AGCCGTGGAGCGGCTGTTGCTAAGGCCCAACAGCTTACGCGGCGGCTTGG 4980
QY 4981 TGCATCCACAGTGGCTGCTC 5000
DB 4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 2
CQ874885 11570 bp DNA linear PAT 27-SBP-2004
LOCUS CQ874885 Sequence 3 from Patent WO2004077060.
DEFINITION CQ874885
ACCESSION CQ874885
VERSION CQ874885.1 GI:52748035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 Diamandis, E.P. and Petraki, C.
Assay for detection of renal cell carcinoma
Patent: WO 2004077060-A 3 10-SBP-2004;
Journal Mount Sinai Hospital (CA)
FEATURES
source 1.11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGAGTGGCTTAGTC 60
DB 1 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGAGTGGCTTAGTC 60
QY 61 TCCCTGCTTAAATGTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 61 TCCCTGCTTAAATGTCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 AAGAAAG 180
DB 121 AAGAAAG 180
QY 181 ACAGAGAGCTGGAGACAGAGGAGACAGAGAGTCAAGAGAGAGAGAGAGAGAG 240
DB 181 ACAGAGAGCTGGAGACAGAGGAGACAGAGAGTCAAGAGAGAGAGAGAGAGAG 240

QY 241 ACACAAATGAGACACAGAGTGTAAAGAAAGAGAGATTAACAGATCCACATACCC 300
DB 241 ACACAAATGAGACACAGAGTGTAAAGAAAGAGAGATTAACAGATCCACATACCC 300
QY 301 AAAAGGAGAGAGACAGAGTTTCAAGGAGGAGTCTATGATCATCTTCTTTTTTTT 360
DB 301 AAAAGGAGAGAGACAGAGTTTCAAGGAGGAGTCTATGATCATCTTCTTTTTTTT 360
QY 361 TTTTTTTTTTTTTTTTGAAGCGAGAGTCTGCTGTCGAGGAGGAGGAGGAGGAG 420
DB 361 TTTTTTTTTTTTTTTTGAAGCGAGAGTCTGCTGTCGAGGAGGAGGAGGAGGAG 420
QY 421 GATCTCGCTCACTGCAAGCTCCGCTCCCGGGTTTCAAGCCATCTCTGCTCAAGCTC 480
DB 421 GATCTCGCTCACTGCAAGCTCCGCTCCCGGGTTTCAAGCCATCTCTGCTCAAGCTC 480
QY 481 CCAAGTACTGGAGCTACAGGCGCCCGCCACTACCGCCGGCTTAATTTTTTGTATTTT 540
DB 481 CCAAGTACTGGAGCTACAGGCGCCCGCCACTACCGCCGGCTTAATTTTTTGTATTTT 540
QY 541 GTAGAGAGCGGGTTTCAACGTTTATAGCGGGAGTGGCTGATCTCTGACCTCGATCC 600
DB 541 GTAGAGAGCGGGTTTCAACGTTTATAGCGGGAGTGGCTGATCTCTGACCTCGATCC 600
QY 601 GCCCGCTCGGCTCCCAAAGTCTGAGATTAACAGGCGTGAACACCGCGCCGAGATC 660
DB 601 GCCCGCTCGGCTCCCAAAGTCTGAGATTAACAGGCGTGAACACCGCGCCGAGATC 660
QY 661 ATCATCTTCTTACATGATGCTGATGTCAGAGTACTTAAAGCCATGAGCTTACCTTTA 720
DB 661 ATCATCTTCTTACATGATGCTGATGTCAGAGTACTTAAAGCCATGAGCTTACCTTTA 720
QY 721 AATATGCAATTTGGGCGAGGACCGTGGCTCATGCTGATTTCCAGACCTTGGAGGC 780
DB 721 AATATGCAATTTGGGCGAGGACCGTGGCTCATGCTGATTTCCAGACCTTGGAGGC 780
QY 781 AGAGGTGGGTGAATCACTTGAAGGCCAGAGATTGAGACCAAGCTGAGCCAACTGTGAA 840
DB 781 AGAGGTGGGTGAATCACTTGAAGGCCAGAGATTGAGACCAAGCTGAGCCAACTGTGAA 840
QY 841 CTCTGCTTTTACTTAAAAAATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
DB 841 CTCTGCTTTTACTTAAAAAATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 901 ACCCTGAATCCAGGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
DB 901 ACCCTGAATCCAGGATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
QY 961 GAGGTTGAGTGGGCGAGATCAATCAACCGCTCCAGCTGGGCGGACAGAGCAAGCT 1020
DB 961 GAGGTTGAGTGGGCGAGATCAATCAACCGCTCCAGCTGGGCGGACAGAGCAAGCT 1020
QY 1021 CTCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
DB 1021 CTCTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
QY 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGTGAATTAAGGAGGAGGAGGAGGAGG 1140
DB 1081 AAAAAAATGCTGTCAACAAATAGAGAGAGTGAATTAAGGAGGAGGAGGAGGAGG 1140
QY 1141 AGAAGTGAAGTATATTGACAAATCTTCAAACTTTTAAAAAGAAAGATCAAGA 1200
DB 1141 AGAAGTGAAGTATATTGACAAATCTTCAAACTTTTAAAAAGAAAGATCAAGA 1200
QY 1201 GGCATTAAGAGACAGAGGAGGAGACAGAGAGACAGAAACCTGTGGCCCAAGAGAGAA 1260
DB 1201 GGCATTAAGAGACAGAGGAGGAGACAGAGAGACAGAAACCTGTGGCCCAAGAGAGAA 1260
QY 1261 ACAAGGCTCTTAAGACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB 1261 ACAAGGCTCTTAAGACAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AAAAAAGAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

1321 AAAAAGCAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGATAGAAAGA 1380
1381 GAGAGAGGGGTGAGAGAGACAGAGATTATTGAGAGAGACTCAGAAAAGATAGCCGAGGGA 1440
1381 GAGAGAGGGGTGAGAGAGACAGAGATTATTGAGAGAGACTCAGAAAAGATAGCCGAGGGA 1440
1441 GAACCAAGAGAGATGGAGAAAGACTGTGAGAAAAAACAGAGACAAGATGGAAAAAGG 1500
1441 GAACCAAGAGAGATGGAGAAAGACTGTGAGAAAAAACAGAGACAAGATGGAAAAAGG 1500
1501 AGTATGAGGGGTAAACAGACAGTGTGGAAATGAGCAAAATGACAGAGAAAGAAACAGAA 1560
1501 AGTATGAGGGGTAAACAGACAGTGTGGAAATGAGCAAAATGACAGAGAAAGAAACAGAA 1560
1561 TCCAGGCGCCAAAGATATGTGACCCAGATTGGTGAAGCCAGATCTTAAAGCTGGGGG 1620
1561 TCCAGGCGCCAAAGATATGTGACCCAGATTGGTGAAGCCAGATCTTAAAGCTGGGGG 1620
1621 AGGCAAGGAAAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCTCCGGGCGAGGAG 1680
1621 AGGCAAGGAAAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCTCCGGGCGAGGAG 1680
1681 GTAGGAGTGAATTCCGGACTGGGTGGGGGGTGTCTGAGGGGTGAGATAGGGGAGCA 1740
1681 GTAGGAGTGAATTCCGGACTGGGTGGGGGGTGTCTGAGGGGTGAGATAGGGGAGCA 1740
1741 GGAAGAGCTATTCTCTAAGGCCCGGATAGGACCTCATTTGCCCGGAAATGTGCCAGGAG 1800
1741 GGAAGAGCTATTCTCTAAGGCCCGGATAGGACCTCATTTGCCCGGAAATGTGCCAGGAG 1800
1741 GGAAGAGCTATTCTCTAAGGCCCGGATAGGACCTCATTTGCCCGGAAATGTGCCAGGAG 1800
1801 CAGTGGTGGTTTAACTCAGAGCCCGGTGCGAGCCAGAGCCAGAGAGAGAGAGAGAGCA 1860
1801 CAGTGGTGGTTTAACTCAGAGCCCGGTGCGAGCCAGAGCCAGAGAGAGAGAGAGAGCA 1860
1801 CAGTGGTGGTTTAACTCAGAGCCCGGTGCGAGCCAGAGCCAGAGAGAGAGAGAGAGCA 1860
1861 AGGCAAGGCTGAGAGAGTCTGCGGCTGAGCTGAGCAAAATCCCCACCCCTACCTGG 1920
1861 AGGCAAGGCTGAGAGAGTCTGCGGCTGAGCTGAGCAAAATCCCCACCCCTACCTGG 1920
1861 AGGCAAGGCTGAGAGAGTCTGCGGCTGAGCTGAGCAAAATCCCCACCCCTACCTGG 1920
1861 AGGCAAGGCTGAGAGAGTCTGCGGCTGAGCTGAGCAAAATCCCCACCCCTACCTGG 1920
1921 GGGAGCAGGGGCAAGTGAAGCTGTGTGAGGGTGTCTCAGCAGGAGGAGAGAGAGTGTCT 1980
1921 GGGAGCAGGGGCAAGTGAAGCTGTGTGAGGGTGTCTCAGCAGGAGGAGAGAGAGTGTCT 1980
1921 GGGAGCAGGGGCAAGTGAAGCTGTGTGAGGGTGTCTCAGCAGGAGGAGAGAGAGTGTCT 1980
1981 GTGCGCTGTGCAACCATTTCTCTGCTCCCTCTGCTGCTGCTGAGAGCTGTCTGAGA 2040
1981 GTGCGCTGTGCAACCATTTCTCTGCTCCCTCTGCTGCTGCTGAGAGCTGTCTGAGA 2040
2041 CTCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCAGTGTGAGCCCTGTC 2100
2041 CTCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCAGTGTGAGCCCTGTC 2100
2101 CTCTGAGTCTCTCTCTAAGTGTGAGGAAATTAAGTGAAGGAGAGAGAGAGTGTGAGA 2160
2101 CTCTGAGTCTCTCTCTAAGTGTGAGGAAATTAAGTGAAGGAGAGAGAGAGTGTGAGA 2160
2101 CTCTGAGTCTCTCTCTAAGTGTGAGGAAATTAAGTGAAGGAGAGAGAGAGTGTGAGA 2160
2161 GCTCCCCGGAATGCGCTGGGCTGCCAACTCTGAGCAATCCCAATCCAGGTGAGAGGCC 2220
2161 GCTCCCCGGAATGCGCTGGGCTGCCAACTCTGAGCAATCCCAATCCAGGTGAGAGGCC 2220
2221 ATGGCTACAG 2280
2221 ATGGCTACAG 2280
2281 CTGGGGGTGACAGGTAACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2281 CTGGGGGTGACAGGTAACTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
2341 CTCTGAGGAG 2400
2341 CTCTGAGGAG 2400
2401 GACCGGATGAATGTAGTCTGTCTGTATCTTGTGAGTGTGAGAGTGTGAGAGAGAGAG 2460
2401 GACCGGATGAATGTAGTCTGTCTGTGTATCTTGTGAGTGTGAGAGTGTGAGAGAGAGAG 2460

2401 GACCGGATGAATGTAGTCTGTCTGTGTATCTTGTGAGTGTGAGAGTGTGAGAGAGAG 2460
2461 GACTGCAAGGT 2520
2461 GACTGCAAGGT 2520
2521 CAGGTGGACCTGACCCCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
2521 CAGGTGGACCTGACCCCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
2521 CAGGTGGACCTGACCCCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
2581 ATTGTGTGTGTCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2581 ATTGTGTGTGTCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2581 ATTGTGTGTGTCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
2641 GT 2700
2641 GT 2700
2701 TCCCTGAGGATCCCGGAGATTGGCGTCAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
2701 TCCCTGAGGATCCCGGAGATTGGCGTCAACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
2761 GT 2820
2761 GT 2820
2821 TTTGT 2880
2821 TTTGT 2880
2821 TTTGT 2880
2881 TCTGT 2940
2881 TCTGT 2940
2941 TTTGT 3000
2941 TTTGT 3000
3001 CAGGT 3060
3001 CAGGT 3060
3061 TGT 3120
3061 TGT 3120
3121 GCTAGGCTGCCCGGAGCGGT 3180
3121 GCTAGGCTGCCCGGAGCGGT 3180
3181 GAGGCAACATGGGCGGT 3240
3181 GAGGCAACATGGGCGGT 3240
3241 GGT 3300
3241 GGT 3300
3241 GGT 3300
3301 GAACCAAGCGGCGGCGGT 3360
3301 GAACCAAGCGGCGGCGGCGGT 3360
3301 GAACCAAGCGGCGGCGGCGGT 3360
3361 GGT 3420
3361 GGT 3420
3421 CTACTAAAAATCAAAAAATTAAGT 3480
3421 CTACTAAAAATCAAAAAATTAAGT 3480
3421 CTACTAAAAATCAAAAAATTAAGT 3480
3481 GGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 3540
3481 GGGAGACTGGGCGAGAAAAATCGCTTGAACCCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGT 3540

QY 3541 TCGCGCATTTGACATCCAGCTGGGCAAGAGCGAAATCCGCTCTCGAAAGAAAAA 3600
Db 3541 TCGCGCATTTGACATCCAGCTGGGCAAGAGCGAAATCCGCTCTCGAAAGAAAAA 3600
QY 3601 GAAAAAAGAGGTAAAGAACAGTAATGGGCAAGAGAGACTGATGATGATGGGCG 3660
Db 3601 GAAAAAAGAGGTAAAGAACAGTAATGGGCAAGAGAGACTGATGATGATGGGCG 3660
QY 3661 ATGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Db 3661 ATGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
QY 3721 GTTTTCATGAGATTAAGAAACCTAGAGCTGCTCCCTCCATGATGAGCCCTTAAG 3780
Db 3721 GTTTTCATGAGATTAAGAAACCTAGAGCTGCTCCCTCCATGATGAGCCCTTAAG 3780
QY 3781 CTGAGCGCTTTCTTCTGCTGCTTGGAAACCTAGCTCCGCTCATGAGCTGACC 3840
Db 3781 CTGAGCGCTTTCTTCTGCTGCTTGGAAACCTAGCTCCGCTCATGAGCTGACC 3840
QY 3841 CCACCTCTTCTCTCAACAGCGCCCTAGGCAAGCTAGTGAACCCGCTTAAGGCA 3900
Db 3841 CCACCTCTTCTCTCAACAGCGCCCTAGGCAAGCTAGTGAACCCGCTTAAGGCA 3900
QY 3901 CACCCCTTTGGGCAAGGCTCAACCCCTATTTCTGAGGATACCTTGAAGCCCTTCA 3960
Db 3901 CACCCCTTTGGGCAAGGCTCAACCCCTATTTCTGAGGATACCTTGAAGCCCTTCA 3960
QY 3961 AGTCAGAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4020
Db 3961 AGTCAGAGCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4020
QY 4021 TGCAGTGGGCTGATCTCGGCTCACTGCAACCTTCTCCAGGTTCAAGTATCTGCT 4080
Db 4021 TGCAGTGGGCTGATCTCGGCTCACTGCAACCTTCTCCAGGTTCAAGTATCTGCT 4080
QY 4081 GCTCCACCTCTGAGTACCTGAGATTAAGGTGCGGCAACGCTGCTGCTGCTGCTG 4140
Db 4081 GCTCCACCTCTGAGTACCTGAGATTAAGGTGCGGCAACGCTGCTGCTGCTGCTG 4140
QY 4141 GTGTCTTTAGTGAAGAGAGGTTTCACTTTGTTGGCAGGCTGCTTCAATCCCAACC 4200
Db 4141 GTGTCTTTAGTGAAGAGAGGTTTCACTTTGTTGGCAGGCTGCTTCAATCCCAACC 4200
QY 4201 TAGAGTATCCGCGCACTCGGCTCCAGAGGCTGGGTTTACAGGCGTGAAGCAAGC 4260
Db 4201 TAGAGTATCCGCGCACTCGGCTCCAGAGGCTGGGTTTACAGGCGTGAAGCAAGC 4260
QY 4261 CCCCAAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAAACCTGACCTG 4320
Db 4261 CCCCAAGCCCAAGTCAAGCTCTTTATAGAGACTCTTAACATGTAAACCTGACCTG 4320
QY 4321 CTAACTAACTCAATTCGAAACCTTCTGCTCCAGGCTTGAACCTCACTGAGGCT 4380
Db 4321 CTAACTAACTCAATTCGAAACCTTCTGCTCCAGGCTTGAACCTCACTGAGGCT 4380
QY 4381 TGAACCCACTTTCTTGAAGCAAGTTCATCCCTAAAGCCTTGGTCTCTCCATCCGAG 4440
Db 4381 TGAACCCACTTTCTTGAAGCAAGTTCATCCCTAAAGCCTTGGTCTCTCCATCCGAG 4440
QY 4441 GCTCCAGCCCAAGCTTGGCACTTACCTGAGCTTGTCAAGAACTCTGATCCCAAT 4500
Db 4441 GCTCCAGCCCAAGCTTGGCACTTACCTGAGCTTGTCAAGAACTCTGATCCCAAT 4500
QY 4501 TTTTACCTCAATGATGTTCTAGCAATTCAGAAATCTGTAGAGTCAAGTGAAGTCA 4560
Db 4501 TTTTACCTCAATGATGTTCTAGCAATTCAGAAATCTGTAGAGTCAAGTGAAGTCA 4560
QY 4561 GTTAACTCACTGAGCTGGGCTGTCTTCTTGAAGCTTGAAGCTTGAAGGTTGCA 4620
Db 4561 GTTAACTCACTGAGCTGGGCTGTCTTCTTGAAGCTTGAAGCTTGAAGGTTGCA 4620

QY 4621 CTCTATTCTCCAGGCGCTGCGCCCTGCGCCCTCAGCATGTCAAGCAACCCCTGAGCT 4680
Db 4621 CTCTATTCTCCAGGCGCTGCGCCCTGCGCCCTCAGCATGTCAAGCAACCCCTGAGCT 4680
QY 4681 GGTCTGGCTCTTGTGATCTGAAACCCCAAGCCCAAGCCCGCTTGAAGCCCGCC 4740
Db 4681 GGTCTGGCTCTTGTGATCTGAAACCCCAAGCCCAAGCCCGCTTGAAGCCCGCC 4740
QY 4741 CAACCAATTTTCGTTCCAGAGCATGTTCTGCGCAACATGATGTTTCTGTGACCA 4800
Db 4741 CAACCAATTTTCGTTCCAGAGCATGTTCTGCGCAACATGATGTTTCTGTGACCA 4800
QY 4801 CCTTAAACCGTGCCTTCTGAGCAACCAAGACTGAGAGCTGAGGCGGAGAAAGC 4860
Db 4801 CCTTAAACCGTGCCTTCTGAGCAACCAAGACTGAGAGCTGAGGCGGAGAAAGC 4860
QY 4861 CCGGTGGATGACAGCAGCAGCCGCTATCATATGATCCGATGCGATATGCAACCC 4920
Db 4861 CCGGTGGATGACAGCAGCAGCCGCTATCATATGATCCGATGCGATATGCAACCC 4920
QY 4921 AGCGTGGAGGCGCGCTGTTGCTAAGGCCCAACAGCTTACTGCGGCGGCTGTTG 4980
Db 4921 AGCGTGGAGGCGCGCTGTTGCTAAGGCCCAACAGCTTACTGCGGCGGCTGTTG 4980
QY 4981 TGCATCAACAGTGGCTGCTC 5000
Db 4981 TGCATCAACAGTGGCTGCTC 5000

RESULT 3
COB74961 11570 bp DNA linear PAT 27-SEP-2004
LOCUS COB74961
DEFINITION Sequence 5 from Patent WO2004075713.
ACCESSION COB74961
VERSION COB74961.1 GI:52748060
KEYWORDS

SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1
AUTHORS Diamandis, E.P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
location/Qualifiers
source 1..11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5000; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGCCAGAGTGAAGCAAGAGAGAGTGAAGCTCCCTGCAAGTGGCTTGAATC 60
Db 1 GGGCCAGAGTGAAGCAAGAGAGAGTGAAGCTCCCTGCAAGTGGCTTGAATC 60
QY 61 TCCCTGCTTAATATGACAGGAGAGGAGCAAGAAACAGGAAAGAGAGGAGGAG 120
Db 61 TCCCTGCTTAATATGACAGGAGAGGAGCAAGAAACAGGAAAGAGAGGAGGAG 120
QY 121 AAGAAAG 180
Db 121 AAGAAAG 180
QY 181 ACAGAGAGCTGGGAGACAGAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGAGACAGAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAG 240
QY 241 ACACAATGAGAGACAGAGGAGTGAAGAAAGAGAGATTAACAGAGTCCAGATCA 300

Db	241	ACACAAATGAGACACAGAGGTGTAAGAAAGAGAGATTAAACAAGTCCAGATACACGC	300
Oy	301	AAAGGGGAGAAAGCAGATTTTCAGGGGGGTGTCATATATCATCTCTTTT	360
Db	301	AAAGGGGAGAAAGCAGATTTTCAGGGGGGTGTCATATATCATCTCTTTT	360
Oy	361	TTTTTTTTTTTTTTGAGACGGAAGTCTCGCTCTGTCCGACAGGCTGAGTGGCGG	420
Db	361	TTTTTTTTTTTTTTGAGACGGAAGTCTCGCTCTGTCCGACAGGCTGAGTGGCGG	420
Oy	421	GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCCATTTCTCGTCAAGCTC	480
Db	421	GATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCCATTTCTCGTCAAGCTC	480
Oy	481	CCAAGTAGCTGGAGCTACAGGCGCCCGGCACTACGCGCGGCTAAATTTTTTTT	540
Db	481	CCAAGTAGCTGGAGCTACAGGCGCCCGGCACTACGCGCGGCTAAATTTTTTTT	540
Oy	541	GTAGAGACGGGGTTTCACTGTTTTAGCCGGATGGCTCGATCTCTGACCTCGTGAATCC	600
Db	541	GTAGAGACGGGGTTTCACTGTTTTAGCCGGATGGCTCGATCTCTGACCTCGTGAATCC	600
Oy	601	GCCGCGCTCGGCTTCCAAAGTGTCTGGGATTAACAGGCGTGAACCAACGCGCCGCAATG	660
Db	601	GCCGCGCTCGGCTTCCAAAGTGTCTGGGATTAACAGGCGTGAACCAACGCGCCGCAATG	660
Oy	661	ATCATCTTCTGAACATATGCTGATGGAACAATGACTTAAGGCATACAGCTTACCTTTTA	720
Db	661	ATCATCTTCTGAACATATGCTGATGGAACAATGACTTAAGGCATACAGCTTACCTTTTA	720
Oy	721	AAATGACAGTTTGGGCGCAGGACACGCTGCTCATGCTGTAAATTCACAGACTTGGAGGC	780
Db	721	AAATGACAGTTTGGGCGCAGGACACGCTGCTCATGCTGTAAATTCACAGACTTGGAGGC	780
Oy	781	AGAGTGGGTGAATCATCTTGAGGCCAGAGATTGAGACCAAGCTTGGCCAAATGCTGAAA	840
Db	781	AGAGTGGGTGAATCATCTTGAGGCCAGAGATTGAGACCAAGCTTGGCCAAATGCTGAAA	840
Oy	841	CTGCTCTTTTCTAATAAAAAAAAAAAAAAAAAAAAAATCAAGCCGGGTGTGTGGGGAC	900
Db	841	CTGCTCTTTTCTAATAAAAAAAAAAAAAAAAAAAAAATCAAGCCGGGTGTGTGGGGAC	900
Oy	901	ACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAAACCTGGAAGCG	960
Db	901	ACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAAACCTGGAAGCG	960
Oy	961	GAGGTGAGTGGGCGGAGATCATCATCCGCTTCAAGCTTGGGCGACAGAGCAAGCT	1020
Db	961	GAGGTGAGTGGGCGGAGATCATCATCCGCTTCAAGCTTGGGCGACAGAGCAAGCT	1020
Oy	1021	CTGCTCTAATAATAATAATAAACAACAAGAGATTGTGTGTAAGTTAGTTATATCT	1080
Db	1021	CTGCTCTAATAATAATAATAAACAACAAGAGATTGTGTGTAAGTTAGTTATATCT	1080
Oy	1081	AAAAAAAAAATGCTGTCAACAAATGAGCAGAGTGAATTAAGGAAATTAATGGGCCA	1140
Db	1081	AAAAAAAAAATGCTGTCAACAAATGAGCAGAGTGAATTAAGGAAATTAATGGGCCA	1140
Oy	1141	AGAACTCTAAGGTATATTGCAAAATCATTCAGAACTTTAAAAAAGAAATCAACA	1200
Db	1141	AGAACTCTAAGGTATATTGCAAAATCATTCAGAACTTTAAAAAAGAAATCAACA	1200
Oy	1201	GGCATGAAAGACAGGAGGAAACAGGAGACAGAAACCTGTGGCCCAAGAGAACAA	1260
Db	1201	GGCATGAAAGACAGGAGGAAACAGGAGACAGAAACCTGTGGCCCAAGAGAACAA	1260
Oy	1261	ACAAAGGCTCTTAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Db	1261	ACAAAGGCTCTTAAGCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Oy	1321	AAAAAGACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGAGAGGATGAAAGA	1380

Db	1321	AAAAAGCAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGCGAGAGGGATTGAAAAGA	1380
Qy	1381	GAGGAGGGGTGAGAGAGACACAGATATTGAGAGAGACTCAGAAAATAGCCGAGGGA	1440
Db	1381	GAGAGAGGGGTGAGAGAGACAGAGATTGAGAGAGACTCAGAAAATAGCCGAGGGA	1440
Qy	1441	GAACCAAGAGAGATGGAAGAGACTTGAGAAAAAACAGAGACAAAAGATGGAAGAAG	1500
Db	1441	GAACCAAGAGAGATGGAAGAGACTTGAGAAAAAACAGAGACAAAAGATGGAAGAAG	1500
Qy	1501	AGTATCGAGGGTGAACAGACAGTGGTGAATGACCAAAATGACAGAAAGAAAGACGAA	1560
Db	1501	AGTATCGAGGGTGAACAGACAGTGGTGAATGACCAAAATGACAGAAAGAAAGACGAA	1560
Qy	1561	TCACAGCGCCCAAGAAATAGTGAACCCAGAGTGTGTGAGAAAGCCAGATCCTTAAGCTGGGG	1620
Db	1561	TCACAGCGCCCAAGAAATAGTGAACCCAGAGTGTGTGAGAAAGCCAGATCCTTAAGCTGGGG	1620
Qy	1621	AGGCAGGGAAGGGGCTGGCTTGCTTCGGAGACCCCTCCCATTTCTCCGGGCCAGGGAG	1680
Db	1621	AGGCAGGGAAGGGGCTGGCTTGCTTCGGAGACCCCTCCCATTTCTCCGGGCCAGGGAG	1680
Qy	1681	GTAAGGAGTGAATTTCCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATGAGGGGGAGCA	1740
Db	1681	GTAAGGAGTGAATTTCCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATGAGGGGGAGCA	1740
Qy	1741	GGAGAGACTATTTGCTAAGGCCCGATTAAGCACTCATTTGCCGGGAAATGTGCCCAAGGAG	1800
Db	1741	GGAGAGACTATTTGCTAAGGCCCGATTAAGCACTCATTTGCCGGGAAATGTGCCCAAGGAG	1800
Qy	1801	CAGTGGGTGATTAACTCAGGCCCGGTGCCAGAGCCCAAGAGAGAGGCAAGTGCACAGGA	1860
Db	1801	CAGTGGGTGATTAACTCAGGCCCGGTGCCAGAGCCCAAGAGAGAGGCAAGTGCACAGGA	1860
Qy	1861	AGGCACAGGCTTGAGAAAGTCTGGCGCTGAAGTGGGAACAAATCCCCACCCCTTAACCTGG	1920
Db	1861	AGGCACAGGCTTGAGAAAGTCTGGCGCTGAAGTGGGAACAAATCCCCACCCCTTAACCTGG	1920
Qy	1921	GGGACAGGGCAAGTGAAGACTGTGTGAGGGTGTCTCAGCAGGCAAGGAGAGAGAGTGTCT	1980
Db	1921	GGGACAGGGCAAGTGAAGACTGTGTGAGGGTGTCTCAGCAGGCAAGGAGAGAGAGTGTCT	1980
Qy	1981	GTGGGTCTTGACCCACAATCTTTCTGTGCCCTCTTGGCCCTGTGTGAGAGCTGTAGA	2040
Db	1981	GTGGGTCTTGACCCACAATCTTTCTGTGCCCTCTTGGCCCTGTGTGAGAGCTGTAGA	2040
Qy	2041	CTCCTATTTCTGAATTTCTAATGTGCTGGGTCTCAGCGCAGTGCCTGATGCTGCTC	2100
Db	2041	CTCCTATTTCTGAATTTCTAATGTGCTGGGTCTCAGCGCAGTGCCTGATGCTGCTC	2100
Qy	2101	CTTGTGGTCTCTCTCTACCTGGGAAATTAAGTAAAGGGAAGGGAAGTGGGTTAAG	2160
Db	2101	CTTGTGGTCTCTCTCTACCTGGGAAATTAAGTAAAGGGAAGGGAAGTGGGTTAAG	2160
Qy	2161	GCTCCCCGGATGCGCTGGGCTTCCCAACCTCTGACAATTCGCCATTCAGGTGACGGCC	2220
Db	2161	GCTCCCCGGATGCGCTGGGCTTCCCAACCTCTGACAATTCGCCATTCAGGTGACGGCC	2220
Qy	2221	ATGCTACAGCAAGACCCCCCTGTGATGTGGGTGTCTGTGTCTGTATCAACAGCTTGTT	2280
Db	2221	ATGCTACAGCAAGACCCCCCTGTGATGTGGGTGTCTGTGTCTGTATCAACAGCTTGTT	2280
Qy	2281	CTGGGGGTCAAGGTAACAGAACTCTGGGGTGGGAAGGTTGTGGAAATGGAGAGACTGT	2340
Db	2281	CTGGGGGTCAAGGTAACAGAACTCTGGGGTGGGAAGGTTGTGGAAATGGAGAGACTGT	2340
Qy	2341	CTCTGCGGCACTAAGCGGCTGTGCTCCCGGGGAACTGTGTGAGGCTGGGCACTGACTCGG	2400
Db	2341	CTCTGCGGCACTAAGCGGCTGTGCTCCCGGGGAACTGTGTGAGGCTGGGCACTGACTCGG	2400
Qy	2401	GACCGGGTGAATGTGAAGTCTGTGTCTGTACTTGTGTGTGTGTGCATTCGTAATGTGGCCCTGT	2460
Db	2401	GACCGGGTGAATGTGAAGTCTGTGTCTGTACTTGTGTGTGTGTGTGCATTCGTAATGTGGCCCTGT	2460

[illegible]

QY	3541	TCGGCCATTGCATCCAGCCTGGGCAACAAGCCGAAACTCCGCTCCGAAAGAAAAA	3600
Db	3541	TCGGCCATTGCATCCAGCCTGGGCAACAAGCCGAAACTCCGCTCCGAAAGAAAAA	3600
QY	3601	GAAGAAAAAGGGTAAAGAACCATGTGATGGGCAACGGAGACATGATGATGGATGGGCG	3660
Db	3601	GAAGAAAAAGGGTAAAGAACCATGTGATGGGCAACGGAGACATGATGATGGATGGGCG	3660
QY	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGAAGAGAGAGATTCACAGGATTGAGAAAGCAT	3720
Db	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGAAGAGAGAGATTCACAGGATTGAGAAAGCAT	3720
QY	3721	GTTTTCATCTGAGAAATTCAGAAACTTGAAGGCTGTCTTCCCTCCATGTGTGGCCCCCTAAG	3780
Db	3721	GTTTTCATCTGAGAAATTCAGAAACTTGAAGGCTGTCTTCCCTCCATGTGTGGCCCCCTAAG	3780
QY	3781	CTGAGCCCTTCTTCTCGATCTGTCTGTGGAAACCTTAGCTCCGACCATGAGCTTAGCC	3840
Db	3781	CTGAGCCCTTCTTCTCGATCTGTCTGTGGAAACCTTAGCTCCGACCATGAGCTTAGCC	3840
QY	3841	CCACCTCTTCTCTCAACCAAGCCCTTAGGCAGACTCTAGTGAACCCCGCTTAAGGCA	3900
Db	3841	CCACCTCTTCTCTCAACCAAGCCCTTAGGCAGACTCTAGTGAACCCCGCTTAAGGCA	3900
QY	3901	CACCCCTTTGGGCGAGGCTCAACCCCTATTCTGTGGGTACCTTTAGAAACCCCTTCAA	3960
Db	3901	CACCCCTTTGGGCGAGGCTCAACCCCTATTCTGTGGGTACCTTTAGAAACCCCTTCAA	3960
QY	3961	AGTCAGAGCTTTTCTTTTCTTTTCTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGGAG	4020
Db	3961	AGTCAGAGCTTTTCTTTTCTTTTCTTTTGGAGACAGTCTGTCTCTCTCCAGGCTGGAG	4020
QY	4021	TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCCTCCAGGTTCAAGTGAATTCGT	4080
Db	4021	TGCAGTGGCGTGAATCTCGGCTCACTGCAACCTCTGCCTCCAGGTTCAAGTGAATTCGT	4080
QY	4081	GCTTCACACTCTCTAGTAGTCTGGGATTAACAGTGTCCGGCCACACCGCTGGCTTAATTTT	4140
Db	4081	GCTTCACACTCTCTAGTAGTCTGGGATTAACAGTGTCCGGCCACACCGCTGGCTTAATTTT	4140
QY	4141	GTCCTTTAGTAGAGACAGGGTTTCACTCTGTGGCCAGGCTGGTCTCAAACTCCCAAC	4200
Db	4141	GTCCTTTAGTAGAGACAGGGTTTCACTCTGTGGCCAGGCTGGTCTCAAACTCCCAAC	4200
QY	4201	TCAGGTGATCCGCCACCTCGGCTCCCAAGTGTCTGGGGTTTACAGGCTGAAGCCACGC	4260
Db	4201	TCAGGTGATCCGCCACCTCGGCTCCCAAGTGTCTGGGGTTTACAGGCTGAAGCCACGC	4260
QY	4261	CCCCAGCCCAAGTCAAGACTCTTTATATAGAGACTCTTAACATGTAACTCTGACCTTGGCC	4320
Db	4261	CCCCAGCCCAAGTCAAGACTCTTTATATAGAGACTCTTAACATGTAACTCTGACCTTGGCC	4320
QY	4321	CTAATCTAATCAATTCGAAACCCCTTCTGTCTGCTCAAGCCCTGACCTCACTGAAGCC	4380
Db	4321	CTAATCTAATCAATTCGAAACCCCTTCTGTCTGCTCAAGCCCTGACCTCACTGAAGCC	4380
QY	4381	TGACCCCACTTCTTGAGACAGTTCGATCCCTTAAGGCTGTCTCCCTCCATCCCCAG	4440
Db	4381	TGACCCCACTTCTTGAGACAGTTCGATCCCTTAAGGCTGTCTCCCTCCATCCCCAG	4440
QY	4441	GCTCCAGCCCCCAAGCTTTGGCACTACCCCTGAGCTGTCTCGAGAAATCTGTAACTCAAT	4500
Db	4441	GCTCCAGCCCCCAAGCTTTGGCACTACCCCTGAGCTGTCTCGAGAAATCTGTAACTCAAT	4500
QY	4501	TTTACCCCTCAATGTAGTTCTAGCAATTCACAGAAATCTGTAGGTCAAGTGTGAAGTCA	4560
Db	4501	TTTACCCCTCAATGTAGTTCTAGCAATTCACAGAAATCTGTAGGTCAAGTGTGAAGTCA	4560
QY	4561	GTAACCCCTACCTGAGGCTGTGTCTGTGAGCTTGAACCTTGAAGCTTGAAGGCTGCA	4620
Db	4561	GTAACCCCTACCTGAGGCTGTGTCTGTGAGCTTGAACCTTGAAGCTTGAAGGCTGCA	4620
QY	4621	CTCTTATTTCTCAGGCGCTGTGCGCTGTGCGCTCAGCATGTACAGACCCACCTTGAAGCT	4680

Db 4621 CTCTTATCTCCAGGCGCTGCTCCCTGCTCCCTGCTGATGATGACACCAACCTCTGAGCT 4680
Qy 4681 GGTCTGCGCTCTTGAAGTCTGAACCAACCCCAAGCCCAAGCCCGCTCTGAGCCCGCC 4740
Db 4681 GGTCTGCGCTCTTGAAGTCTGAACCAACCCCAAGCCCAAGCCCGCTCTGAGCCCGCC 4740
Qy 4741 CAACCCATTTTCGCTTCCGAGAGATGTTCTCCGCAACATGATGTTTCTCTGACCAAC 4800
Db 4741 CAACCCATTTTCGCTTCCGAGAGATGTTCTCCGCAACATGATGTTTCTCTGACCAAC 4800
Qy 4801 CCTCTAACACCGTCCCTCTGAGAGACCAAGCACTGAGAGCTGGGAGGAGAGAGC 4860
Db 4801 CCTCTAACACCGTCCCTCTGAGAGACCAAGCACTGAGAGCTGGGAGGAGAGAGC 4860
Qy 4861 CCCGCTGAGTGAACAGAGAGCCGATCATCAATGATCGACTGCGATATGACACCC 4920
Db 4861 CCCGCTGAGTGAACAGAGAGCCGATCATCAATGATCGACTGCGATATGACACCC 4920
Qy 4921 AGCGTGGCAGCGCGCTGTTGTAAGCCCAACAGCTCTACTGCGGGGCGTGTGG 4980
Db 4921 AGCGTGGCAGCGCGCTGTTGTAAGCCCAACAGCTCTACTGCGGGGCGTGTGG 4980
Qy 4981 TGGATCCACAGTGGCTGCTC 5000
Db 4981 TGGATCCACAGTGGCTGCTC 5000

RESULT 4
AF135028 11570 bp DNA linear PRI 27-JUN-2000

LOCUS Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
AF135028
VERSION AF135028.1 GI:4589282

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 11570)
Mammalia, Eutheria, Primates, Catarrhini, Homnidae, Homo.
AUTHORS Yousef, G.M., and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
J Biol. Chem. 274 (53), 37511-37516 (1999)

JOURNAL MEDLINE 20076408
PUBMED 10608802

REFERENCE 2 (bases 1 to 11570)
Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A. and Obiezu, C.V.
TITLE The new human kallikrein gene family: implications in
carcinogenesis
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)

JOURNAL MEDLINE 21121728
PUBMED 10675891

REFERENCE 3 (bases 1 to 11570)
Yousef, G.M., Luo, L.Y. and Diamandis, E.P.
TITLE Direct Submission
Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada

FEATURES
source 1. 11570
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101..2131,2210..2293,4762..5023,5763..6019,
6105..6238,11092..11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221..2293,4762..5023,5763..6019,6105..6238,
11092..11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"

/protein_id="AAD26429.1"
/db_xref="GI:4589283"
translation="MATTARPMMVLCALITALLGVTEHVLANDVSCRPSTYPS
GSDQDLAGAGEDARSDDSSRTINSDDMHTQPMQALILPNDLYCAVALVHPQW
LRLPAHCRKRVPRVRLGHYSLSVPSVGQMFQGVSIHPHVSLSHSLMDLILKLN
RLRIPKDYRPIVSHSPASAGTKLVSMGTTKSPQVHFPLYCLNLSVLSQKCE
DAYRQIDDTMFCAKGRDSCQGSQSPVVCNSGLQGLVSGVDYPCARPNRPGVYT
NICKPTKNIQETIQANS"

Query Match 100.0%; Score 5000; DB 9; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGCCCAAGTGAAGCAAGAGAGAGTGAAGAGTCCCTCTGCAAGTGGCTTGAATC 60
Db 1 GGGCCCAAGTGAAGCAAGAGAGAGTGAAGAGTCCCTCTGCAAGTGGCTTGAATC 60
Qy 61 TCCCTGCTGCTAAATGACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 TCCCTGCTGCTAAATGACAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Qy 121 AAGAAAG 180
Db 121 AAGAAAG 180
Qy 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 ACAGAGAGCTGGGACACAGGAGACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 300
Qy 301 AAAGGGGAG 360
Db 301 AAAGGGGAG 360
Qy 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 420
Db 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 420
Qy 421 GATCTGGGCTCATGAGAGCTGCGCTCCCGGGTTCAGGCAATTCCTGCGCTCAGGCTC 480
Db 421 GATCTGGGCTCATGAGAGCTGCGCTCCCGGGTTCAGGCAATTCCTGCGCTCAGGCTC 480
Qy 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCCGCTAACTTTTGTATTTT 540
Db 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCCGCTAACTTTTGTATTTT 540
Qy 541 GTAGAGACGAGGAGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAGGTT 600
Db 541 GTAGAGACGAGGAGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAGGTTTCAAGGTT 600
Qy 601 GCGCGCTGCGCTCCCAAGGCTGGGATTAACAGGCTGAGACACGCGCGCGCGCATG 660
Db 601 GCGCGCTGCGCTCCCAAGGCTGGGATTAACAGGCTGAGACACGCGCGCGCGCATG 660
Qy 661 ATCATCTTCTGACTAGTGAAGTGAACAAGTACCTAAAGCCATCAGACTCTACCTTT 720
Db 661 ATCATCTTCTGACTAGTGAAGTGAACAAGTACCTAAAGCCATCAGACTCTACCTTT 720
Qy 721 AATATGAGTTTGGGCAAGGCAAGGCTCATGCTGTAATTCAGCACTTTGGAGGC 780
Db 721 AATATGAGTTTGGGCAAGGCAAGGCTCATGCTGTAATTCAGCACTTTGGAGGC 780
Qy 781 AAGAGGAGGAGATCATCTTGAAGGCGAGGTTTGAACAAGCTGGGCAACATGGTAA 840
Db 781 AAGAGGAGGAGATCATCTTGAAGGCGAGGTTTGAACAAGCTGGGCAACATGGTAA 840
Qy 841 CTCGTCTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900
Db 841 CTCGTCTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900

QY 901 ACCTGTAATCCAGATATGCTGAGGCTGAGCAGAGAGTCACTTGAACCTTGAGGCG 960
Db 901 ACCTGTAATCCAGATATGCTGAGGCTGAGCAGAGAGTCACTTGAACCTTGAGGCG 960
QY 961 GAGGTTGCAGTGGGCGGAGATCAATCAACCGCTTCCAGCTGGGCGACAGAGAACT 1020
Db 961 GAGGTTGCAGTGGGCGGAGATCAATCAACCGCTTCCAGCTGGGCGACAGAGAACT 1020
QY 1021 CTGTCTCAATTAATAATAATAACAACAGAGCTTGTGTCTCTTAAGTAAATCT 1080
Db 1021 CTGTCTCAATTAATAATAATAACAACAGAGCTTGTGTCTCTTAAGTAAATCT 1080
QY 1081 AAAAAAAAAATGCTGTCAACAAATAGAGCAGAAAGTAAATAAAGAAAAATAAATGGGCGCA 1140
Db 1081 AAAAAAAAAATGCTGTCAACAAATAGAGCAGAAAGTAAATAAAGAAAAATAAATGGGCGCA 1140
QY 1141 AGAATCTTAAGTATATTTTGAACAAATCATTCAGAACCTTTAAAAAGAAATCAACA 1200
Db 1141 AGAATCTTAAGTATATTTTGAACAAATCATTCAGAACCTTTAAAAAGAAATCAACA 1200
QY 1201 GGCATGAAGAGAGAGGAGAGACAGGAGACAGAAACACTGTGGCCCAAGAGAAACAA 1260
Db 1201 GGCATGAAGAGAGAGGAGAGACAGGAGACAGAAACACTGTGGCCCAAGAGAAACAA 1260
QY 1261 ACAGGCTCCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAGGCTCCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 AAAAAACAG 1380
Db 1321 AAAAAACAG 1380
QY 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTTGAAGAGACTCGAAGAAATAGCCGAGGGA 1440
Db 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTTGAAGAGACTCGAAGAAATAGCCGAGGGA 1440
QY 1441 GAAACACAG 1500
Db 1441 GAAACACAG 1500
QY 1501 AGTATGAGAGGTGAACAGACAGTGTGTGAATGAGCAAAATGACAGAGAGAGAGAGAG 1560
Db 1501 AGTATGAGAGGTGAACAGACAGTGTGTGAATGAGCAAAATGACAGAGAGAGAGAGAG 1560
QY 1561 TCCAGGCGCCAGAAATAGTGAACCCAGAGTTGTGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TCCAGGCGCCAGAAATAGTGAACCCAGAGTTGTGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 AGGCAAGGAGAGGAGCTGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 AGGCAAGGAGAGGAGCTGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GTTGGAGAGTGAACATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GTTGGAGAGTGAACATTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 GAGAGAGCTATTGTAAGGCGGAGTGAAGCACTATTGCGCGGAAATGTGCGCCAGGGAG 1800
Db 1741 GAGAGAGCTATTGTAAGGCGGAGTGAAGCACTATTGCGCGGAAATGTGCGCCAGGGAG 1800
QY 1801 CAGTGGGTGTTTAACTCAGAGCCGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 CAGTGGGTGTTTAACTCAGAGCCGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
QY 1861 AGGCAACAGGCTGAAGAGTGTGCGGCTGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 1920
Db 1861 AGGCAACAGGCTGAAGAGTGTGCGGCTGAGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 1920
QY 1921 GGGAGAGGGCAATGAG 1980
Db 1921 GGGAGAGGGCAATGAG 1980

QY 1981 GTGCGTCTGACACCCACATCTTTCTGTGCTCCCTCCCTGCTGTCTGAGAGCTGTAGA 2040
Db 1981 GTGCGTCTGACACCCACATCTTTCTGTGCTCCCTCCCTGCTGTCTGAGAGCTGTAGA 2040
QY 2041 CTCTATCTTCTGAATTTCTATGTGCTGAGGCTCTCAGGCAATGCCATGTGTGCCCTG 2100
Db 2041 CTCTATCTTCTGAATTTCTATGTGCTGAGGCTCTCAGGCAATGCCATGTGTGCCCTG 2100
QY 2101 CTGTGTGTTCTCTCTCACTGAGGAAATTAAGGTAGGGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 CTGTGTGTTCTCTCTCACTGAGGAAATTAAGGTAGGGAGAGAGAGAGAGAGAGAGAG 2160
QY 2161 GCTCCCGAGATGCGTGGGCTTCCACCCCTTGAATTTCCCATCGAGTGCAGCGGC 2220
Db 2161 GCTCCCGAGATGCGTGGGCTTCCACCCCTTGAATTTCCCATCGAGTGCAGCGGC 2220
QY 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTGTGTCTGTATCAACCTTGTCT 2280
Db 2221 ATGGCTACAGCAAGACCCCTGAGATGTGGGTCTGTGTCTGTATCAACCTTGTCT 2280
QY 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGAGGAGGTTGGGATTTGGAGAGACTGT 2340
Db 2281 CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGAGAGGAGGTTGGGATTTGGAGAGACTGT 2340
QY 2341 CTCTCGGCACTAAGCGCTGTCTCCCTGGGAACTGTGTGAGCTCTGGGCACTGCTCGG 2400
Db 2341 CTCTCGGCACTAAGCGCTGTCTCCCTGGGAACTGTGTGAGCTCTGGGCACTGCTCGG 2400
QY 2401 GACCGGGTGAATGTAGTCTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 GACCGGGTGAATGTAGTCTCTGTCTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
QY 2461 GACTGCAACGGGT 2520
Db 2461 GACTGCAACGGGT 2520
QY 2521 CAGGTGCACTGACCCCTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2521 CAGGTGCACTGACCCCTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
QY 2581 ATTTGT 2640
Db 2581 ATTTGT 2640
QY 2641 GTTTTGT 2700
Db 2641 GTTTTGT 2700
QY 2701 TCCCTGAAGGTCCCGGAGATTGCGTGAACAAAAGTGTATCAATGGAAGCTGTGACT 2760
Db 2701 TCCCTGAAGGTCCCGGAGATTGCGTGAACAAAAGTGTATCAATGGAAGCTGTGACT 2760
QY 2761 GTTGT 2820
Db 2761 GTTGT 2820
QY 2821 TTTTGTGAACGCTGTGACTACTGAAGCTCTGTGTGAAGGAGCTGTATGTGAAGCTGTGTG 2880
Db 2821 TTTTGTGAACGCTGTGACTACTGAAGCTCTGTGTGAAGGAGCTGTATGTGAAGCTGTGTG 2880
QY 2881 TCTGTGTGAAGCGGTGTAAATGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db 2881 TCTGTGTGAAGCGGTGTAAATGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
QY 2941 TTTGT 3000
Db 2941 TTTGT 3000
QY 3001 CAGGTGACTGACTTGAAGT 3060
Db 3001 CAGGTGACTGACTTGAAGT 3060
QY 3061 TGTGACACACCGCATCTGT 3120

Db 3061 TGTGACACACGCGATCTGTGCGTGGACCTGAGACATGTGTGATGAGGTGTGTGATCC 3120
Qy 3121 GCTAGGCTGCCCGGAGCGGTGTGTACTGTGAGACAGAGCTGTATGTATGCTGTGACCTGTG 3180
Db 3121 GCTAGGCTGCCCGGAGCGGTGTGTACTGTGAGACAGAGCTGTATGTATGCTGTGACCTGTG 3180
Qy 3181 GAGGCAACATGAGGCGGTGTGTGAGACATGTGTGCTGTGTGCTGTATGCTGTGTGTG 3240
Db 3181 GAGGCAACATGAGGCGGTGTGTGAGACATGTGTGCTGTGTGCTGTATGCTGTGTGTG 3240
Qy 3241 GCTG 3300
Db 3241 GCTG 3300
Qy 3301 GAGGCAACGCGGCGGCGGTGTGTGACCTGTGTATCCAGCCCTTTGGAGGCGGAGGCA 3360
Db 3301 GAGGCAACGCGGCGGCGGTGTGTGACCTGTGTATCCAGCCCTTTGGAGGCGGAGGCA 3360
Qy 3361 GGGGATACCTGAGGTGCGGAGATGAGGCGAGCTGTGACCAATGTGAGAACCCGCT 3420
Db 3361 GGGGATACCTGAGGTGCGGAGATGAGGCGAGCTGTGACCAATGTGAGAACCCGCT 3420
Qy 3421 CTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3480
Db 3421 CTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3480
Qy 3481 GGGAGACTGGGGGAGAAATTCGTTGAAACCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTG 3540
Db 3481 GGGAGACTGGGGGAGAAATTCGTTGAAACCCGAGAGTGTGTGTGTGTGTGTGTGTGTGTG 3540
Qy 3541 TGTGCGCATGTGACTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Db 3541 TGTGCGCATGTGACTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3600
Qy 3601 GAAAG 3660
Db 3601 GAAAG 3660
Qy 3661 ATGATGT 3720
Db 3661 ATGATGT 3720
Qy 3721 GTTTTATCTGAAATTCAGAAATCTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3780
Db 3721 GTTTTATCTGAAATTCAGAAATCTAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3780
Qy 3781 CTGAGCCCTTCTTCTGT 3840
Db 3781 CTGAGCCCTTCTTCTGT 3840
Qy 3841 CCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3900
Db 3841 CCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3900
Qy 3901 CACCCCTTGTGGGCGAGGCTCAACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Db 3901 CACCCCTTGTGGGCGAGGCTCAACCCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3960
Qy 3961 AGTCAGAGCTTTTGT 4020
Db 3961 AGTCAGAGCTTTTGT 4020
Qy 4021 TGCAGTGGCGT 4080
Db 4021 TGCAGTGGCGT 4080
Qy 4081 GCTTCCACTTCTGT 4140
Db 4081 GCTTCCACTTCTGT 4140
Qy 4141 GTGTCTTGT 4200
Db 4141 GTGTCTTGT 4200

Db 4141 GTGTCTTGT 4200
Qy 4201 TGCAGT 4260
Db 4201 TGCAGT 4260
Qy 4261 CCCCAGCCCAAGT 4320
Db 4261 CCCCAGCCCAAGT 4320
Qy 4321 CTATTAAGT 4380
Db 4321 CTATTAAGT 4380
Qy 4381 TGAACCCCACTTGT 4440
Db 4381 TGAACCCCACTTGT 4440
Qy 4441 GCTCAGCCCCCACTGT 4500
Db 4441 GCTCAGCCCCCACTGT 4500
Qy 4501 TTTACCTGT 4560
Db 4501 TTTACCTGT 4560
Qy 4561 GTTACCTGT 4620
Db 4561 GTTACCTGT 4620
Qy 4621 CTCTTATTTCTGT 4680
Db 4621 CTCTTATTTCTGT 4680
Qy 4681 GGT 4740
Db 4681 GGT 4740
Qy 4741 CAACCAATTTTGT 4800
Db 4741 CAACCAATTTTGT 4800
Qy 4801 CCTCTAACCGGCTGT 4860
Db 4801 CCTCTAACCGGCTGT 4860
Qy 4861 CCGGT 4920
Db 4861 CCGGT 4920
Qy 4921 AGCGGT 4980
Db 4921 AGCGGT 4980
Qy 4981 TGCATCAAGT 5000
Db 4981 TGCATCAAGT 5000

RESULT 5
AC011483/c 107487 bp DNA linear PRI 01-JUL-2002
LOCUS Homo sapiens chromosome 19 clone CTB-147c22, complete sequence.
DEFINITION AC011483
AC011483
VERSION AC011483.7 GI:21637461
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submision

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 107487)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 1, 2002 this sequence version replaced gi:14971176.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.hgsc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.6.
NOTE: Shatter libraries failed to resolve dinucleotide repeat.
Unsure number of repeat copies 64998-65494. Forced join 65015.
Location/Qualifiers
source 1..107487
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTB-147C2"
misc_feature 64998..65494
/note="NOTE: Shatter libraries failed to resolve dinucleotide repeat. Unsure number of repeat copies 64998-65494. Forced join 65015."
ORIGIN
Query Match 99.8%; Score 4989; DB 9; Length 107487;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGCCAGAGTGAAGGCAAGAGAGATTGAGAGCTCCCTGCAAAATGGCTTGAGTC 60
DB 60220 GGGCCAGAGTGAAGGCAAGAGAGATTGAGAGCTCCCTGCAAAATGGCTTGAGTC 60161
QY 61 TCCCTGCTTAATATCAGAGAGAGGAGGCAAGAAACAGGGA- GAGGAAGGGGTGGG 119
DB 60160 TCCCTGCTTAATATCAGAGAGAGGAGGAGGCAAGAAACAGGGAAGGAGGTGGG 60101
QY 120 GAAGAAAGAGAGAGAGAGAGAGAGATTAACAACACTACAGAAACACAGAGAGAA 179
DB 60100 GAAGAAAGAGAGAGAGAGAGAGAGATTAACAACACTACAGAAACACAGAGAGAA 60041
QY 180 CACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAAAGAGAAATAGAGAA 239
DB 60040 CACAGAGAGCTGGGACACAGGAGACACAGAGTCAGAGAGAAAGAGAAATAGAGAA 59981
QY 240 GACACAATATGAGACACAGAGGTGTAAAGAAAGAGATTTAACAGAGTCCCAAGTACAG 299
DB 59980 GACACAATATGAGACACAGAGGTGTAAAGAAAGAGATTTAACAGAGTCCCAAGTACAG 59921
QY 300 CAAAGGGGACAGAGCAGTTTTCAGGGGTGTATATCATCTCTTTTCTTTTCTTTT 359
DB 59920 CAAAGGGGACAGAGCAGTTTTCAGGGGTGTATATCATCTCTTTTCTTTTCTTTT 59861
QY 360 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 419
DB 59860 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 59801
QY 420 GGAATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCAATCTCTGCTCAAGCT 479
DB 59800 GGAATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCAAGCAATCTCTGCTCAAGCT 59741
QY 480 CCCAAGTAGCTGGAGCTACAGGGGCGGCACTAGGCGCGGCTAATTTTGTATTTT 539
DB 59740 CCCAAGTAGCTGGAGCTACAGGGGCGGCACTAGGCGCGGCTAATTTTGTATTTT 59681

QY 540 AGTAGAGACGGGGTTTACCCGTTTACCGGGAGTGGCTGATCTCTGACCTCGTATC 599
DB 59680 AGTAGAGACGGGGTTTACCCGTTTACCGGGAGTGGCTGATCTCTGACCTCGTATC 59621
QY 600 CCGCCGCTCGGCTTCCCAAATGCTGGGATTACAGGGGTGAGCCACCGCCCGCCAT 659
DB 59620 CCGCCGCTCGGCTTCCCAAATGCTGGGATTACAGGGGTGAGCCACCGCCCGCCAT 59561
QY 660 GATCATCTTCTGACATAGCTGATGAGCAAGATACCTTAAGGCATCAGCTACCTTT 719
DB 59560 GATCATCTTCTGACATAGCTGATGAGCAAGATACCTTAAGGCATCAGCTACCTTT 59501
QY 720 AATATGACATTGAGGACAGGACCGTGCTCATCTGTAAATTCACAGACTTTGGAGG 779
DB 59500 AATATGACATTGAGGACAGGACCGTGCTCATCTGTAAATTCACAGACTTTGGAGG 59441
QY 780 CAGAGTGGGTGAATCATCTTGAGCGCAGAGATTGAGACACGCTGGCCAAATGTGAA 839
DB 59440 CAGAGTGGGTGAATCATCTTGAGCGCAGAGATTGAGACACGCTGGCCAAATGTGAA 59381
QY 840 ACTGCTCTTTCTTAAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 899
DB 59380 ACTGCTCTTTCTTAAAAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 59321
QY 900 CACTGTATATCCACAGCTATGCTGAGGCTGAGGACAGAGGTCACTTGAACCTGAGAG 959
DB 59320 CACTGTATATCCACAGCTATGCTGAGGCTGAGGACAGAGGTCACTTGAACCTGAGAG 59261
QY 960 GGAAGTGTGAGTGGGCGGAGATCATATCCGCTCCAGGCTGGGCGCAAGAGCAAG 1019
DB 59260 GGAAGTGTGAGTGGGCGGAGATCATATCCGCTCCAGGCTGGGCGCAAGAGCAAG 59201
QY 1020 TCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1079
DB 59200 TCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 59141
QY 1080 TAAAAAATATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGG 1139
DB 59140 TAAAAAATATGCTGTCAACAAATAGACAGAGTGAATTAAGAAAAATTAATGGG 59081
QY 1140 AAGAACTTAAGTATATTTTGAACAAATCATTTTAAAAAAGAAATCAACG 1199
DB 59080 AAGAACTTAAGTATATTTTGAACAAATCATTTTAAAAAAGAAATCAACG 59021
QY 1200 AGGCATTAAGAGACAGGGAGGAAACAGGAGACAGAAACCTGTGGCCCAAGAGAA 1259
DB 59020 AGGCATTAAGAGACAGGGAGGAAACAGGAGACAGAAACCTGTGGCCCAAGAGAA 58961
QY 1260 AACAAAGCTCTTAAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1319
DB 58960 AACAAAGCTCTTAAGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 58901
QY 1320 GAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAGAGGGATGAA 1379
DB 58900 GAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAGAGGGATGAA 58841
QY 1380 AGAGAGAGGGGTGAGAGAGACAGAGATTTGAAGAGACTCAGAAAAGTAGCCGAGG 1439
DB 58840 AGAGAGAGGGGTGAGAGAGACAGAGATTTGAAGAGACTCAGAAAAGTAGCCGAGG 58781
QY 1440 AGAACACAGAGAGATGAGAGAGACTGTGAGAAAAAACAAGAGCAAAAGATGAGAA 1499
DB 58780 AGAACACAGAGAGATGAGAGAGACTGTGAGAAAAAACAAGAGCAAAAGATGAGAA 58721
QY 1500 GAGTATCGAGGGTGAACAGACAGTGTGTGAATGAGCAAAATGACAGAGAAAGCAAG 1559
DB 58720 GAGTATCGAGGGTGAACAGACAGTGTGTGAATGAGCAAAATGACAGAGAAAGCAAG 58661
QY 1560 ATCCAGGGGCAAGAAATGTGACCCAGAGTTGATGAGAGCAAGATCCTTAAGGCTGG 1619
DB 58660 ATCCAGGGGCAAGAAATGTGACCCAGAGTTGATGAGAGCAAGATCCTTAAGGCTGG 58601
QY 1620 GAGGACGAGAAAGGGGCTGGCTGGCTCCGAGAACCTCTCCCATTTCTCGGGCAAGG 1679

Db 5860 GAGGACAGGGAAGGGGCTGGCTGGCTCCGAGACCCCTCCCACTTCTCCGGGCGACAGGA 58541
Qy 1680 GGTAGGGAGTGACATTCGCGACTGGGTGGGGGGGTCTCTGGGGGTGAGATAGGGGGAGC 1739
Db 58540 GGTAGGGAGTGACATTCGCGACTGGGTGGGGGGGTCTCTGGGGGTGAGATAGGGGGAGC 58481
Qy 1740 AGGAGGAGCTATTGCTAAGGCGCGATAGGACCTCATTTGCCCGGGAAATGAGCCCAAGGA 1799
Db 58480 AGGAGGAGCTATTGCTAAGGCGCGATAGGACCTCATTTGCCCGGGAAATGAGCCCAAGGA 58421
Qy 1800 GCAGTGGGTGTTATTACTAGAGCCCGGTGCCAGAGCCCAAGAGAGAGAGAGTGGCAGG 1859
Db 58420 GCAGTGGGTGTTATTACTAGAGCCCGGTGCCAGAGCCCAAGAGAGAGAGTGGCAGG 58361
Qy 1860 AAGGCACAGGCTTGAGAGTCTGGGCTGAGCTGGAGCAAAATCCCACTCCCTACCTG 1919
Db 58360 AAGGCACAGGCTTGAGAGTCTGGGCTGAGCTGGAGCAAAATCCCACTCCCTACCTG 58301
Qy 1920 GGGGACAGGGCAATGAGACTGGTGAAGGTGGCTCAGCAGGCGAGAGAGAGAGTGC 1979
Db 58300 GGGGACAGGGCAATGAGACTGGTGAAGGTGGCTCAGCAGGCGAGAGAGAGAGTGC 58241
Qy 1980 TGTGCGTCTGCAACCACTCTTCTCTGCTCCCTCCCTGCTGCTGAGAGCTGCTAG 2039
Db 58240 TGTGCGTCTGCAACCACTCTTCTCTGCTCCCTCCCTGCTGCTGAGAGCTGCTAG 58181
Qy 2040 ACTCTATCTTTGTAATTTATAGTGCCTGGGCTCAGCGCAGTGCAGATGGCCCT 2099
Db 58180 ACTCTATCTTTGTAATTTATAGTGCCTGGGCTCAGCGCAGTGCAGATGGCCCT 58121
Qy 2100 CCTTGGGTCTCTCTCACTGGGGAAATAGTAGAGGAGAGAGAGAGAGTGGTTAG 2159
Db 58120 CCTTGGGTCTCTCTCACTGGGGAAATAGTAGAGGAGAGAGAGAGTGGTTAG 58061
Qy 2160 GGGTCCCCGGATGCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGGTGACGGG 2219
Db 58060 GGGTCCCCGGATGCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGGTGACGGG 58001
Qy 2220 CATGGCTACAGCAAGAACCCCTCGATGAGGTGCTCTGTCTGATCAACGCTTGT 2279
Db 58000 CATGGCTACAGCAAGAACCCCTCGATGAGGTGCTCTGTCTGATCAACGCTTGT 57941
Qy 2280 TCTGGGGGTCAAGGTAAACAAGAACTCTGGGTGGAGGTTGGATTTGGAGAGACTG 2339
Db 57940 TCTGGGGGTCAAGGTAAACAAGAACTCTGGGTGGAGGTTGGATTTGGAGAGACTG 57881
Qy 2340 TCTTGGGCACTAGAGCGCTGCTCCCTGGGGAAGTGTGAGACTGGGATGACTCG 2399
Db 57880 TCTTGGGCACTAGAGCGCTGCTCCCTGGGGAAGTGTGAGACTGGGATGACTCG 57821
Qy 2400 GAGACCGGAGTAAATGAGTCTGTCTGTACTTGTGTGTGCGATGTAATGAGCCCTG 2459
Db 57820 GAGACCGGAGTAAATGAGTCTGTCTGTACTTGTGTGTGCGATGTAATGAGCCCTG 57761
Qy 2460 TGACTGCAAGGTGTGTGTGGGAGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCG 2519
Db 57760 TGACTGCAAGGTGTGTGTGGGAGAGGGGATGCTTTTCCCATATCAGGTGACTGTGCG 57701
Qy 2520 GCAAGTGGCACTGACCCCTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTAA 2579
Db 57700 GCAAGTGGCACTGACCCCTTGAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTAA 57641
Qy 2580 GATTGTGTGTGTCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCTG 2639
Db 57640 GATTGTGTGTGTCTCAACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTCTG 57581
Qy 2640 TGT 2699
Db 57580 TGT 57521
Qy 2700 GTCCCTGAGGTCCCGGAGTTGCGTGAACAATAAGTGTGATCACAATGAAAGCTGTGAC 2759

Db 57520 GTCCCTGAGGTCCCGGAGTTGCGTGAACAATAAGTGTGATCACAATGAAAGCTGTGAC 57461
Qy 2760 TGT 2819
Db 57460 TGT 57401
Qy 2820 ATTTGTGACCGGTGTGACTACCTGAACTCTGTGTAGGGGTGACTGTATGTGACTGTGT 2879
Db 57400 ATTTGTGACCGGTGTGACTACCTGAACTCTGTGTAGGGGTGACTGTATGTGACTGTGT 57341
Qy 2880 GTCTGTGTGAGGCGGTGTAAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGAG 2939
Db 57340 GTCTGTGTGAGGCGGTGTAAATGCTACTGTATGTGTGTGTGTGTGTGTGTGTGTGTGAG 57281
Qy 2940 TTTTCTGTCTGTGCTGAGAGGATAGAGGGGTGACAGGGGTAGCTATCTTGTGGAAATGGGT 2999
Db 57280 TTTTCTGTCTGTGCTGAGAGGATAGAGGGGTGACAGGGGTAGCTATCTTGTGGAAATGGGT 57221
Qy 3000 CCAAGTGACTGACTTGCAGTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3059
Db 57220 CCAAGTGACTGACTTGCAGTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 57161
Qy 3060 CTGTGACACACGGCATCTGT 3119
Db 57160 CTGTGACACACGGCATCTGT 57101
Qy 3120 CGCTAAGCTGTCCCGGAGGCTGT 3179
Db 57100 CGCTAAGCTGTCCCGGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 57041
Qy 3180 GAGAGCAACATGGGCGGT 3239
Db 57040 GAGAGCAACATGGGCGGT 56981
Qy 3240 CGGT 3299
Db 56980 CGGT 56921
Qy 3300 AGAACAGAGCTGGGCGGT 3359
Db 56920 AGAACAGAGCTGGGCGGT 56861
Qy 3360 AGGCGGATCACTGAGGTGGGAGATGAGAGCCAGCTGACCAATGAGAAACCCCTG 3419
Db 56860 AGGCGGATCACTGAGGTGGGAGATGAGAGCCAGCTGACCAATGAGAAACCCCTG 56801
Qy 3420 TCTACTAATAATAAATAAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3479
Db 56800 TCTACTAATAATAAATAAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 56741
Qy 3480 CCGGAGACTGGGCGAATAATGCTTTGAAACCGGAGGTGTGAGGTTCGGGTGACCGAG 3539
Db 56740 CCGGAGACTGGGCGAATAATGCTTTGAAACCGGAGGTGTGAGGTTCGGGTGACCGAG 56681
Qy 3540 ATGGCGCATTTGACCTCCAGCTGGGCAACAAGAGGAAATCCGCTCGAAAGAAATAA 3599
Db 56680 ATGGCGCATTTGACCTCCAGCTGGGCAACAAGAGGAAATCCGCTCGAAAGAAATAA 56621
Qy 3600 AGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3659
Db 56620 AGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 56561
Qy 3660 CATGCAATGATCTGTAGGTCTGT 3719
Db 56560 CATGCAATGATCTGTAGGTCTGT 56501
Qy 3720 TGTTCATCTGAGATTCAGAAACCAAGGCTGTCTTCCCTCATGTGTGTGTGTGTGTGTGT 3779
Db 56500 TGTTCATCTGAGATTCAGAAACCAAGGCTGTCTTCCCTCATGTGTGTGTGTGTGTGTGT 56441
Qy 3780 GCTGAGGCTTCTTCTGT 3839
Db 56440 GCTGAGGCTTCTTCTGT 56381

QY 3840 CCCACCTCTTCTCTCAACCAAGCCCTAGGCGCAGACTCTAGTGAAGCCCGGCTTAAGGCC 3899
DB 56380 CCCACTCTTCTCTCAACCAAGCCCTAGGCGCAGACTCTAGTGAAGCCCGGCTTAAGGCC 56321
QY 3900 ACAACCCCTTGGGCGAGGCTCAACCCCTATTCTGTGGGTACTTCTTAAGAACCCCTTCA 3959
DB 56320 ACAACCCCTTGGGCGAGGCTCAACCCCTATTCTGTGGGTACTTCTTAAGAACCCCTTCA 56261
QY 3960 AAGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTCCCGAGGCTGA 4019
DB 56260 AAGTCAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGTCTGTCTCTCCCGAGGCTGA 56201
QY 4020 GTGCACTGGCGTATCTCGGCTCACTGCAACTCTGCTCTCCAGTTCAAGTATCTCG 4079
DB 56200 GTGCACTGGCGTATCTCGGCTCACTGCAACTCTGCTCTCCAGTTCAAGTATCTCG 56141
QY 4080 TGCCTCACTCTCTGAGTACGTTGAGATTACAGGTGCGCGCACACGCTGCTAAATTTT 4139
DB 56140 TGCCTCACTCTCTGAGTACGTTGAGATTACAGGTGCGCGCACACGCTGCTAAATTTT 56081
QY 4140 TGTGCTTTAGTAGACAGGGTTTCACTTGTGGCGAGGCTGTCTCAAACTCCCAAC 4199
DB 56080 TGTGCTTTAGTAGACAGGGTTTCACTTGTGGCGAGGCTGTCTCAAACTCCCAAC 56021
QY 4200 CTGAGGTGATCCGCCCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAACACCG 4259
DB 56020 CTGAGGTGATCCGCCCACTCGGCTCCAGAGTGTGGGTTACAGGCGTGAACACCG 55961
QY 4260 CCCCCAGCCCAAGTCAAGACTCTTTTAAAGAGACTTAAATCTGAACCTGACCTGCG 4319
DB 55960 CCCCCAGCCCAAGTCAAGACTCTTTTAAAGAGACTTAAATCTGAACCTGACCTGCG 55901
QY 4320 CCTAAGTGAATTCCTAAACCCCTCTGCTCCAGCCCTCAACCCCACTCACTGAGGC 4379
DB 55900 CCTAAGTGAATTCCTAAACCCCTCTGCTCCAGCCCTCAACCCCACTCACTGAGGC 55841
QY 4380 CTGACCCCACTTCTTGAAGCAAGTTCATCCCTAAAGCCCTGCTCCCTCCATCCCA 4439
DB 55840 CTGACCCCACTTCTTGAAGCAAGTTCATCCCTAAAGCCCTGCTCCCTCCATCCCA 55781
QY 4440 GGGTCAGGCCCCCAAGCTTTGGCACTACCCCTGAGCTTGTCCAGAAATCTGTACCA 4499
DB 55780 GGGTCAGGCCCCCAAGCTTTGGCACTACCCCTGAGCTTGTGTACAGAAATCTGTACCA 55721
QY 4500 TTTTACCTCACTAGTATCTAGCCAAATCTCGAAGATCTGAGGTCCAGTTAGAGTCC 4559
DB 55720 TTTTACCTCACTAGTATCTAGCCAAATCTCGAAGATCTGAGGTCCAGTTAGAGTCC 55661
QY 4560 AGTAACTCACTAGGCTGTGGCTCTGCTTGTAGCTTGAAGCTTGAAGAGTGC 4619
DB 55660 AGTAACTCACTAGGCTGTGGCTCTGCTTGTAGCTTGAAGCTTGAAGAGTGC 55601
QY 4620 ACTCTTATCTCAAGGCTGCGCCCTGCGCCCTCAAGATGTCAACCCACTTATGAC 4679
DB 55600 ACTCTTATCTCAAGGCTGCGCCCTGCGCCCTCAAGATGTCAACCCACTTATGAC 55541
QY 4680 TGTGCTGGCTCTTGAAGTGAACCAACCCCAAGCCCGCTGTGAGCCCGC 4739
DB 55540 TGTGCTGGCTCTTGAAGTGAACCAACCCCAAGCCCGCTGTGAGCCCGC 55481
QY 4740 CCAACCATTTTCCGTTCCAGAGCAATGTTCTGCCAACAATGATGTTTCTGTGACAC 4799
DB 55480 CCAACCATTTTCCGTTCCAGAGCAATGTTCTGCCAACAATGATGTTTCTGTGACAC 55421
QY 4800 CCTCTTAACAACGCTGCTCTTGGAGCAACCAAGAACCTTGGAGCTGCGGCGGAGAAC 4859
DB 55420 CCTCTTAACAACGCTGCTCTTGGAGCAACCAAGAACCTTGGAGCTGCGGCGGAGAAC 55361
QY 4860 GCGCGGTGAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 4919
DB 55360 GCGCGGTGAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 55301

QY 4920 CAGCGGTGAGCAGCGCGCTGTGGTAGAGCCCAACAGCTCTACTGCGGGCGGTGTG 4979
DB 55300 CAGCGGTGAGCAGCGCGCTGTGGTAGAGCCCAACAGCTCTACTGCGGGCGGTGTG 55241
QY 4980 GTGCATCCACAGTGGCTGCTC 5000
DB 55240 GTGCATCCACAGTGGCTGCTC 55220

RESULT 6
AF243527/c 230000 bp DNA linear PRI 21-NOV-2000
LOCUS Homo sapiens serine protease gene cluster, complete sequence.
DEFINITION AF243527
ACCESSION AF243527
VERSION AF243527.1 GI:11244757
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Sequencing and expression analysis of the serine protease gene
cluster located in chromosome 19q13 region
JOURNAL Gene 257 (1), 119-130 (2000)
MEDLINE 20510030
PUBMED 11054574
REFERENCE
AUTHORS 2 (bases 1 to 230000)
Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,
Moss, P., Paepel, B. and Wang, K.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,
Bothell, WA 98021, USA
FEATURES
source location/Qualifiers
1..230000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13"
/complement(<6963..>11517)
/gene="KLK1"
complement(join(<6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/product="renal kallikrein"
complement(join(6963..7118,7668..7804,7923..8212,
9482..9641,11472..11517))
/gene="KLK1"
/note="serine protease"
/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCIALSLGTAAPPIQSHIVGMECEHSPQMAALYH
PSTFOGGLVHROWLTAALHCISDNYOLALGRHLFPDENTAOAVHVESPPHGFN
MSLENNHROADEVSHDMLRLNPEADTITDAVKVIELPTEPEVUSTCLASMGK
IEPENSPFDIQCVDLKITLPDECKKAHVQVYTPMFCVGHLSGKDTCVDSGGL
MCDGVLAQVTSAGYVPCGTPNKPVSVAVLSVKXIEPTIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
/product="ACO protease"
complement(join(13552..13704,14377..14530,14678..14917,
15416..15569,19204..>19246))
/note="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MMILTLTSLFLASTPAADGDKLHGDECAPIHSQPMQVALYERGR
FNCGASLISPHVNLAAHQSRFMRVLRGEHLNRKRDPEQLRTTSRYIPIHRYEARS
HNNDIMLRIVQPARLNPQVRPAVLPTRCPHGEACVSVSGMWLVLSPVSLPDLTLC

OY		61	TCCCTCGCTTAAATTCAGGAGAGGGAGGCAGAAAGACAGGGA-GAGGAAGGGTGGG	119
Dd		142488	TCCCTCGCTTAAATTCAGGAGAGGGAGGCAGAAAGACAGGGAAGGGGTGGG	142429
OY		120	GAAAGAAAGAGAGAGAGAGACAAGAATTAACTAACAGAAACAAGAGAAACA	179
Dd		142428	GAAAGAAAGAGAGAGAGAGACAAGATTAACTAACAGAAACAAGAGAAACA	142369
OY		180	CACAGAGAGCTTGGGACACAGGGAACACAGAGTCAAGAGAAAAGAAAGATTAGAAAA	239
Dd		142368	CACAGAGAGCTTGGGACACAGGGAACACAGAGTCAAGAGAAAAGAAAGATTAGAAAA	142309
OY		240	GACACAAATGGAGACAACAAGGTATAAAGAAAGAGATTAAACAGTCCCAGATCACG	299
Dd		142308	GACACAAATGGAGACAACAAGGTATAAAGAAAGAGATTAAACAGTCCCAGATCACG	142249
OY		300	CAAAGGGGACAGAAACACATTTTCAGGGTGTGTCTATGATCATCTCTTTTTTTTTT	359
Dd		142248	CAAAGGGGACAGAAACACATTTTCAGGGTGTGTCTATGATCATCTCTTTTTTTTTT	142189
OY		360	TTTTTTTTTTTTTTTTGAGACGGAGTCTGCTGTGTGCCAGGCTGAGGTGAGTGGC	419
Dd		142188	TTTTTTTTTTTTTTTTGAGACGGAGTCTGCTGTGTGCCAGGCTGAGGTGAGTGGC	142129
OY		420	GGATCTCGGCTCACAGCAAGCTCCGCGCTCCCGGTTCAAGCCATTTCCCTGACTGCT	479
Dd		142128	GGATCTCGGCTCACAGCAAGCTCCGCGCTCCCGGTTCAAGCCATTTCCCTGACTGCT	142069
OY		480	CCCAAGTAGCTGGACCTACAGGCGCCGACCTACGCCCGGCTAATTTTTTGATTTT	539
Dd		142068	CCCAAGTAGCTGGACCTACAGGCGCCGACCTACGCCCGGCTAATTTTTTGATTTT	142009
OY		540	AGTAGAGACGGGGTTTCACTGTTTAAGCCGGATGGCTTGATCTCTGACTCTGTGATC	599
Dd		142008	AGTAGAGACGGGGTTTCACTGTTTAAGCCGGATGGCTTGATCTCTGACTCTGTGATC	141949
OY		600	CGCCGCGCTCGGCTCCCAAAGTGTGGGATTTACAGGCTGAGGCCACCOCGCCGCGCAT	659
Dd		141948	CGCCGCGCTCGGCTCCCAAAGTGTGGGATTTACAGGCTGAGGCCACCOCGCCGCGCAT	141889
OY		660	GATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCCATCAGACTCACTCTT	719
Dd		141888	GATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCCATCAGACTCACTCTT	141829
OY		720	AAATATGCAGTTTGGGCGACGCGTGTCTATGCTGTAAATTCAGACACTTTGGAGG	779
Dd		141828	AAATATGCAGTTTGGGCGACGCGTGTCTATGCTGTAAATTCAGACACTTTGGAGG	141769
OY		780	CAGAGGTGGGTGAATCACTTGAAGCCAGAGATTGAGACAGGCTGGCCAACATGTGTGA	839
Dd		141768	CAGAGGTGGGTGAATCACTTGAAGCCAGAGATTGAGACAGGCTGGCCAACATGTGTGA	141709
OY		840	ACTGTGTCTTTACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTGCTGGAGCA	899
Dd		141708	ACTGTGTCTTTACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTGCTGGAGCA	141649
OY		900	CACCTGTAAATCCAGCTTATGCTGGAGCTTGAGGACAGAGTCACTTGAACCTTGAAGC	959
Dd		141648	CACCTGTAAATCCAGCTTATGCTGGAGCTTGAGGACAGAGTCACTTGAACCTTGAAGC	141589
OY		960	GGAGGTGACAGTGGGCGAGATCAACATCACCGGCTCCAGGCTGGGGCGACAGGCAAGAC	1019
Dd		141588	GGAGGTGACAGTGGGCGAGATCAACATCACCGGCTCCAGGCTGGGGCGACAGGCAAGAC	141529
OY		1020	TCTGTCTCAATTAATTAATTAACAACAACAGACAGTGTGTGTACTTATGTTATATC	1079
Dd		141528	TCTGTCTCAATTAATTAATTAACAACAACAGACAGTGTGTGTACTTATGTTATATC	141469
OY		1080	TMAAAAAAAAAATGCTGTCAACAATAGACAGAGTGAATTAAGAAAAATTAATGGGCC	1139
Dd		141468	TMAAAAAAAAAATGCTGTCAACAATAGACAGAGTGAATTAAGAAAAATTAATGGGCC	141409

QY	1140	AAGAACTGTAAGGTATATTTGCAATCAATCTCAGAAACCTTTAAAAAAGAAACATCACG	1199
Db	141408	AAGAATCTTAAGGTATATTTGCAATCAATCAATCTTTAAAAAAGAAATCACG	141349
QY	1200	AGGCATTAGAAAGACAGGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA	1259
Db	141348	AGGCATTAGAAAGACAGGGAGGAAACAGGAGACAGAAACAACCTGTGGCCCAAGAGAACAA	141289
QY	1260	AAACAAGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGTGAAGTGAAGACAGACAGA	1319
Db	141288	AAACAAGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGTGAAGTGAAGACAGACAGA	141229
QY	1320	GAAAAAGACAGAGAGAGAGACAGACAGAGAGACAGAGAGGGGAGGGGATGGAAG	1379
Db	141228	GAAAAAGACAGAGAGAGAGACAGACAGAGAGACAGAGAGGGGAGGGGATGGAAG	141169
QY	1380	AGAGAGAGGGGTGAGAGAGACACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGG	1439
Db	141168	AGAGAGAGGGGTGAGAGAGACACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGG	141109
QY	1440	AGAACACAGAGAGATGGAAGAGAACTCTGAGAAAAAACAGAGACAAAGATGAAAGAG	1499
Db	141108	AGAACACAGAGAGATGGAAGAGAACTCTGAGAAAAAACAGAGACAAAGATGAAAGAG	141049
QY	1500	CAGTATCGAAGGTGGAACAGACAGTGTGTGAATGACAAATATCCAGAGAGAAAGCAGGA	1559
Db	141048	CAGTATCGAAGGTGGAACAGACAGTGTGTGAATGACAAATATCCAGAGAGAAAGCAGGA	140989
QY	1560	ATCCAGGCGCCAAAGATAGTGAACCCAGAGTTGGTGAAGACAGATCCTTAAAGCTGGGG	1619
Db	140988	ATCCAGGCGCCAAAGATAGTGAACCCAGAGTTGGTGAAGACAGATCCTTAAAGCTGGGG	140929
QY	1620	GAGGCAAGGAAAGGGCTGGCTGGCTTCGAGAACCCCTCCCATTTCTCCGGGCAAGGA	1679
Db	140928	GAGGCAAGGAAAGGGCTGGCTGGCTTCGAGAACCCCTCCCATTTCTCCGGGCAAGGA	140869
QY	1680	GGTAGAGAGTGAATTTCCGGAATGGGTGGGGGTGTCTGTGGGGGTGAGATAGGGGGAGC	1739
Db	140868	GGTAGAGAGTGAATTTCCGGAATGGGTGGGGGTGTCTGTGGGGGTGAGATAGGGGGAGC	140809
QY	1740	AGGAGAGACTATTTGTTAAAGGCCCAATAGGACCTCATTTGCCGGGAAATGTGCCCAAGGA	1799
Db	140808	AGGAGAGACTATTTGTTAAAGGCCCAATAGGACCTCATTTGCCGGGAAATGTGCCCAAGGA	140749
QY	1800	GCAGTGGGTGTTAATCTCAGGCCCGGTGCCCAGAGCCCAAGAGAGAGCACTGTGCCAG	1859
Db	140748	GCAGTGGGTGTTAATCTCAGGCCCGGTGCCCAGAGCCCAAGAGAGAGCACTGTGTGCCAG	140689
QY	1860	AAAGGACAGGCTTGAGAAATGTGCGCGCTGAAGTGTGGAGCAAAATCCCCACCCCTTAACCTG	1919
Db	140688	AAAGGACAGGCTTGAGAAATGTGCGCGCTGAAGTGTGGAGCAAAATCCCCACCCCTTAACCTG	140629
QY	1920	GGGGACAGGGCAAGAGAGACCTGTGAGAGGTGGCTCAGCAGGACAGGAAAGAGAGAGGTGTC	1979
Db	140628	GGGGACAGGGCAAGAGAGACCTGTGAGAGGTGGCTCAGCAGGACAGGAAAGAGAGAGGTGTC	140589
QY	1980	TGTGAGTCTGTGACCCACATCTTTCTGTGCCCTCTTGGCCCTGTCTTGAGAGCTGCTAG	2039
Db	140568	TGTGAGTCTGTGACCCACATCTTTCTGTGCCCTCTTGGCCCTGTCTTGAGAGCTGCTAG	140509
QY	2040	ACTCTTAATCTTGAATTTATATAGTGCCTGGGTCTCAGCGCAGTGCCTGATGTGCCCCGT	2099
Db	140508	ACTCTTAATCTTGAATTTATATAGTGCCTGGGTCTCAGCGCAGTGCCTGATGTGCCCCGT	140449
QY	2100	CCTTGTGTTCTCTCTAATCTGAGGGAATTAAGTGAAGGGAAGGGAAGGGGATTAAG	2159
Db	140448	CCTTGTGTTCTCTCTAATCTGAGGGAATTAAGTGAAGGGAAGGGAAGGGGATTAAG	140389
QY	2160	GGCTCCCGGATCGCTGGGCTCTCCCAACCCCTGTACAATTCGCCATCCAGGTGCAGCGCG	2219
Db	140388	GGCTCCCGGATCGCTGGGCTCTCCCAACCCCTGTACAATTCGCCATCCAGGTGCAGCGCG	140329
QY	2220	CATGGCTACAGCAGAACCCCGGTGAGTGGGTGCTGTGTGCTGTGATCACAGCCTTGCT	2279

Dh	140328	CATGGCTACAGCAAGACCCCTTGGAATGTGGGTCTCTGTCTCTGATCAACGCTTGTCT	140269
Qy	2280	TTGGGGGATCAAGATTAACAGAACTTGAGGGTGGAGGTTGTGGATTGGAGAGACTG	2339
Dh	140268	TTTGGGGGATCAAGATTAACAGAACTTGAGGGTGGAGGTTGTGGATTGGAGAGACTG	140209
Qy	2340	TTCTTGGGGGATCAAGATTAACAGAACTTGAGGGTGGAGGTTGTGGATTGGAGAGACTG	2339
Dh	140208	TTCTTGGGGGATCAAGATTAACAGAACTTGAGGGTGGAGGTTGTGGATTGGAGAGACTG	140149
Qy	2400	GGACGGGGATTAAGTCTCTGTCTGATCTTGATGTTGGTGGATGATGAGGCTG	2459
Dh	140148	GGACGGGGATTAAGTCTCTGTCTGATCTTGATGTTGGTGGATGATGAGGCTG	140089
Qy	2460	TGACTGCCACGGTGTGTGTGGGGAGGGAGTGCCTTTTCCATATACAGTGACTGTGCG	2519
Dh	140088	TGACTGCCACGGTGTGTGTGGGGAGGGAGTGCCTTTTCCATATACAGTGACTGTGCG	140029
Qy	2520	GCAAGTGGCACTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2579
Dh	140028	GCAAGTGGCACTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	139969
Qy	2580	GATTGT	2639
Dh	139968	GATTGT	139909
Qy	2640	TGT	2699
Dh	139908	TGT	139849
Qy	2700	GTCTCTAAGGTCTCCGGGATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2759
Dh	139848	GTCTCTAAGGTCTCCGGGATTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	139789
Qy	2760	TGT	2819
Dh	139788	TGT	139729
Qy	2820	ATTTGTGACCGGT	2879
Dh	139728	ATTTGTGACCGGT	139669
Qy	2880	GTCTGT	2939
Dh	139668	GTCTGT	139609
Qy	2940	TTTCTGT	2999
Dh	139608	TTTCTGT	139549
Qy	3000	CCAGGTGATCTGT	3059
Dh	139548	CCAGGTGATCTGT	139489
Qy	3060	CTGT	3119
Dh	139488	CTGT	139429
Qy	3120	CGT	3179
Dh	139428	CGT	139369
Qy	3180	GGAGGCAACATGAGCGTGTCTGAGAACTGCGTGTGTGTGTGTGTGTGTGTGTGTGT	3239
Dh	139368	GGAGGCAACATGAGCGTGTCTGAGAACTGCGTGTGTGTGTGTGTGTGTGTGTGTGT	139309
Qy	3240	CGCGT	3299
Dh	139308	CGCGT	139249
Qy	3300	AGAACAGGCGGCGGCGGT	3359
Dh	139248	AGAACAGGCGGCGGCGGT	139189
Qy	3360	AGCGGATTCATCTGAGGTGGGAGATTCGAGGCCAGCTGTGACCAATGAGAGAACTCCG	3419
Dh	139188	AGCGGATTCATCTGAGGTGGGAGATTCGAGGCCAGCTGTGACCAATGAGAGAACTCCG	139129
Qy	3420	TTTACTTAAATTAACAAATTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3479
Dh	139128	TTTACTTAAATTAACAAATTAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	139069
Qy	3480	CGGAGACTGGGGCAGAAAAATCGCTTGAACCCGGAGGTGAGAGTTGCGGTGACCGAG	3539
Dh	139068	CGGAGACTGGGGCAGAAAAATCGCTTGAACCCGGAGGTGAGAGTTGCGGTGACCGAG	139009
Qy	3540	ATGCGGCAATTGCACTTCAGCTGGGCAACAAGAGGAACTCCGTCTCCAAAGAAAAA	3599
Dh	139008	ATGCGGCAATTGCACTTCAGCTGGGCAACAAGAGGAACTCCGTCTCCAAAGAAAAA	138949
Qy	3600	AGAAAAAAGGATTAAGCAAGTGAATGGGCAAGGAGAGCTGATGATGAGTGGG	3659
Dh	138948	AGAAAAAAGGATTAAGCAAGTGAATGGGCAAGGAGAGCTGATGATGAGTGGG	138889
Qy	3660	CATGATGTAGTCTGTAGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3719
Dh	138888	CATGATGTAGTCTGTAGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	138829
Qy	3720	TGT	3779
Dh	138828	TGT	138769
Qy	3780	GCTGAGGCTTCTTCTTCTGCTGCTGCTTGTGGAACCTAGCTCCGCCATGAGCTGTAC	3839
Dh	138768	GCTGAGGCTTCTTCTTCTGCTGCTGCTTGTGGAACCTAGCTCCGCCATGAGCTGTAC	138709
Qy	3840	CCCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	3899
Dh	138708	CCCACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	138649
Qy	3900	ACACCCCTTGGGCGAGGCTTCAACCCCTTATTTCTGTGGGTACTTCTTGAACCCCTTCA	3959
Dh	138648	ACACCCCTTGGGCGAGGCTTCAACCCCTTATTTCTGTGGGTACTTCTTGAACCCCTTCA	138589
Qy	3960	AACTGAGGCTTTTTTTTTTTTTTTTTTTTTTGTGAGACATCTGTCTCTCCAGGCTGGA	4019
Dh	138588	AACTGAGGCTTTTTTTTTTTTTTTTTTTTTTGTGAGACATCTGTCTCTCCAGGCTGGA	138529
Qy	4020	GTGCAGTGGGTATCTCGGCTCACTGCAACTCTGTGCTCCAGGTTCAAGTATTTCTG	4079
Dh	138528	GTGCAGTGGGTATCTCGGCTCACTGCAACTCTGTGCTCCAGGTTCAAGTATTTCTG	138469
Qy	4080	TGCTTCACTCTCTGAGTACTGGGATTAAGGTGGCGGCAACAGCTGTGCTAATTTT	4139
Dh	138468	TGCTTCACTCTCTGAGTACTGGGATTAAGGTGGCGGCAACAGCTGTGCTAATTTT	138409
Qy	4140	TGTGTCTTTTGTGAGACAGGGTTTCACTTGTGTGGCAGGCTGTCTCAACTCCCAAC	4199
Dh	138408	TGTGTCTTTTGTGAGACAGGGTTTCACTTGTGTGGCAGGCTGTCTCAACTCCCAAC	138349
Qy	4200	CTGAGGTATCCGCCCACTCGGCTCCAGAGTGTGTGGGTTTACAGGCTGTGAGCCACG	4259
Dh	138348	CTGAGGTATCCGCCCACTCGGCTCCAGAGTGTGTGGGTTTACAGGCTGTGAGCCACG	138289
Qy	4260	CCCCCAGCCAAAGTCAAGCTCTTTTATGAGAGCTTAAATGATGTAACCTGTACCTGTG	4319
Dh	138288	CCCCCAGCCAAAGTCAAGCTCTTTTATGAGAGCTTAAATGATGTAACCTGTACCTGTG	138229
Qy	4320	CCTTAACTGATCAATTCAAACCCCTTCCGCTCCAGGCTGTGACCCCACTGACTGAGG	4379
Dh	138228	CCTTAACTGATCAATTCAAACCCCTTCCGCTCCAGGCTGTGACCCCACTGACTGAGG	138169
Qy	4380	CTGACCCCACTTCTTGAACCAAGTTCATCTTAAAGCCTGTGTCTTCCATCCCA	4439
Dh	138168	CTGACCCCACTTCTTGAACCAAGTTCATCTTAAAGCCTGTGTCTTCCATCCCA	138109

QY	4440	GGCTCCAGCCCCCACAAGCTTTGGACATACCCCGACCTTGTCACAGAAATCTGTACCCAA	4499
Db	138108	GGCTCCAGCCCCCACAAGCTTTGGACATACCCCGACCTTGTCACAGAAATCTGTACCCAA	138049
QY	4500	TTTAAACCTCACATGATGTTCTAGCCCAATTCAGAAATCTGTAGAGTCCAGTTAGAGTCC	4559
Db	138048	TTTAAACCTCACATGATGTTCTAGCCCAATTCAGAAATCTGTAGAGTCCAGTTAGAGTCC	1379889
QY	4560	AGTAACCCCTACCTGAGCCCTGGGCTCTGTCTTGAAGTTAGAGCTTGAGAGAGTGC	4619
Db	137988	AGTAACCCCTACCTGAGCCCTGGGCTCTGTCTTGAAGTTAGAGCTTGAGAGAGTGC	1379299
QY	4620	ACTCTTAATCTCCAGAGCCCTGGCCCTCGCCCTCAGCATGTACAGACCAACCCCTTAGC	4679
Db	137928	ACTCTTAATCTCCAGAGCCCTGGCCCTCGCCCTCAGCATGTACAGACCAACCCCTTAGC	1378658
QY	4680	TGCTCTGAGCTCTTGTAGTCTGAAACCCACCCCAAGCCCGCTCTGAGCCCGC	4739
Db	137868	TGCTCTGAGCTCTTGTAGTCTGAAACCCACCCCAAGCCCGCTCTGAGCCCGC	137809
QY	4740	CCAAACCCATTTCCGTTCCCAAGCAATGTTCCGCAATGATATTTTCCTGTGACAC	4799
Db	137808	CCAAACCCATTTCCGTTCCCAAGCAATGTTCCGCAATGATATTTTCCTGTGACAC	137749
QY	4800	CCCTTAAACACCGCTGCTCTGTGGAGCAACACGAGCTTGAGCTGGAGCGGGAGAGAC	4859
Db	137748	CCCTTAAACACCGCTGCTCTGTGGAGCAACACGAGCTTGAGCTGGAGCGGGAGAGAC	1376889
QY	4860	GCCCGGTGGATGACAGACAGACGCCGATCATATGATGATCGATCGATATGACACAC	4919
Db	137688	GCCCGGTGGATGACAGACAGACGCCGATCATATGATGATCGATCGATATGACACAC	1376299
QY	4920	CAGCGGTGGACAGCGCGCTGTGCTAAAGGCCCAACACACTCTGATCGCGGGGAGGTG	4979
Db	137628	CAGCGGTGGACAGCGCGCTGTGCTAAAGGCCCAACACACTCTGATCGCGGGGAGGTG	1375699
QY	4980	GTCATCCACACAGTGGCTGCTC 5000	
Db	137568	GTCATCCACACAGTGGCTGCTC 137548	
RESULT 7			
AC027602/c			
LOCUS	217346 bp	DNA	linear
DEFINITION	Homo sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT		
ACCESSION	AC027602.4	GI:11178143	
VERSION	AC027602.4	GI:11178143	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Baeten, V., Bedalov, F., Boguski, M., Boulton, A., Boudreau, B., Brown, A., Burkett, G., Campilioni, A., Castle, A., Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeRubeis, K., Dewar, K., Diaz, J.S., Dodgson, S., Domingo, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardy, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J., Iley, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margus, N., McCarthy, M., McEwan, P., McGarr, A., McKernan, K., McPherson, R., Meldrum, J., Meneses, L., Mihov, T., Mizrahi, C., Miska, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,		

TITLE
JOURNAL
REFERENCE
AUTHORS

O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanic, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliou, H., Vriel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 217346)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Beda, F., Bonnaville, L., Boutkhalil, B., Brown, A., Burkett, G., Campiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaidyna, S., Ginde, S., Goyette, M., Graham, J., Grand-Pierre, N., Hagos, B., Hearford, A., Horton, L., Ilium, I., Johnson, R., Jones, C., Kam, L., Karatas, A., Labocque, K., Lamaras, R., Landers, T., Lehotzky, J., Levine, R., Liu, C., Liu, G., Macdonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menush, L., Mhova, T., Mlenka, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N., Pisanic, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J., Vassiliou, H., Vriel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 15, 2000 this sequence version replaced gi:11136831.
 All repeats were identified using RepeatMasker:
 Smit, A.P.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9166
 Center clone name: 795_B6
 ----- Summary Statistics
 Sequencing vector: M13: M77815; 31% of reads
 Sequencing vector: Plasmid; n/a; 69% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 210748 bases at least Q40
 Consensus quality: 21655 bases at least Q30
 Consensus quality: 215058 bases at least Q20
 Insert size: 19400; agarose-fp
 Insert size: 216246; sum-of-contigs
 Quality coverage: 11.9 in Q20 bases; agarose-fp
 Quality coverage: 10.7 in Q20.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 8150 8149: contig of 8149 bp in length
 8150 8249: gap of 100 bp
 8250 9592: contig of 1343 bp in length
 9593 9692: gap of 100 bp
 9693 10733: contig of 1041 bp in length
 10734 10833: gap of 100 bp
 10834 13519: contig of 2686 bp in length

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L9166
Center clone name: 795_B.6
-----
Summary Statistics
Sequencing vector: M13, M77815, 31% of reads
Sequencing vector: Plasmid; n/a; 63% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 21655 bases at least Q30
Consensus quality: 215058 bases at least Q20
Insert size: 194000; agarose-fp
Insert size: 216246; sum-of-consigs
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.
* NOTE: This is a "working draft" sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
8149: contig of 8149 bp in length
1
8150
8249: gap of 100 bp
8250
9592: contig of 1343 bp in length
9593
9692: gap of 100 bp
9693
10733: contig of 1041 bp in length
10734
10833: gap of 100 bp
10834
13519: contig of 2686 bp in length

```


FEATURES	
*	13520: gap of 100 bp
*	13620: contig of 3891 bp in length
*	17511: gap of 100 bp
*	17611: contig of 6932 bp in length
*	24602: gap of 100 bp
*	24702: gap of 100 bp
*	35434: contig of 10732 bp in length
*	35435: gap of 100 bp
*	35533: gap of 100 bp
*	124474: contig of 88940 bp in length
*	124574: gap of 100 bp
*	124575: contig of 10090 bp in length
*	134665: gap of 100 bp
*	134764: gap of 100 bp
*	162343: contig of 27579 bp in length
*	162344: gap of 100 bp
*	162345: gap of 100 bp
*	208817: contig of 46674 bp in length
*	209017: gap of 100 bp
*	208918: gap of 100 bp
*	217346: contig of 8339 bp in length.
	Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP1-795B6"
/clone_1fb="RPCI-11 Human Male BAC"
1. .8159
/note="assembly_fragment
clone_end:SP6
vector_side:left"
8250. .9592
/note="assembly_fragment"
9693. .10733
/note="assembly_fragment"
10834. .13519
/note="assembly_fragment"
13620. .17510
/note="assembly_fragment"
17611. .24602
/note="assembly_fragment"
24703. .35434
/note="assembly_fragment"
35535. .124474
/note="assembly_fragment"
124575. .1194664
/note="assembly_fragment"
134765. .162343
/note="assembly_fragment"
162444. .208917
/note="assembly_fragment"
209018. .217346
/note="assembly_fragment
clone_end:T7
vector_side:right"
```

ORIGIN

	84.9%	Score 4243.4	DB 2	Length 217346
Query March	Similarity 99.0%	Pred. No. 0		
Best Local	Conservative	Mismatches	Indels	Gaps
Matches 4335	0	11	35	5
QY	654	GGCCATGATCATCTTCTTGACTATGCTGATGTGACAGTAACTTAAGCCATCAGACTCTA	713	
Db	203789	GTCATATGATCATCTTCTTGACTATGCTGATGTGACAGTAACTTAAGCCATCAGACTCTA	203730	
QY	714	CCCTTTAATAATGACATTGGGCGCAGGCAACCGTGCTACCTGTAAATTCAGCACTTT	773	
Db	203729	CCCTTTAATAATGACATTGGGCGCAGGCAACCGTGCTACCTGTAAATTCAGCACTTT	203670	
QY	774	GGGAGGCGAGAGGTGGGTGAAATCACTTGAGCCAGAGGTTGAGACCAAGCTGGCCAAAT	833	
Db	203669	GGGAGGCGAGAGGTGGGTGAAATCACTTGAGCCAGAGGTTGAGACCAAGCTGGCCAAAT	203610	
QY	834	GGTGAACCTGTCTTTAATAAAAAAAAAAAAAAAAAAATCAGCGGGGTGTGTT	893	

Db	203609	GGTGAACCTCTGCTCTTACT-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAATCAGCGGGGTGTGCT	203551
QY	894	GGGGCACACTGTAAATCCCACTATGCTGAGGCTGAGGCACAGAGTCACTTGAACCTT	953
Db	203550	GGGGGACACCTGTAAATCCCACTATGCTGAGGCTGAGGCACAGAGTCACTTGAACCTT	203491
QY	954	GGAGCGGAGGTTGAGATGTTGGGCCCGAGATCAATCACCGCCCTCCAGCCTTGGGCGACAGG	1013
Db	203490	GGAGCGGAGGTTGAGATGTTGGGCCCGAGATCAATCACCGCCCTCCAGCCTTGGGCGACAGG	203431
QY	1014	CAAGACTCTGCTCAAAATAAATAAACAACAGACAGTTGTGTAACCTTAGT	1073
Db	203430	CAAGACTCTGCTCAAAATAAATAAACAACAGACAGTTGTGTAACCTTAGT	203371
QY	1074	TATATCTAAAAAAAAAATGCTGTCAACAAATAGACAGAGTGAAATAAAGGAAATATA	1133
Db	203370	TATATCTAAAAAAAAAATGCTGTCAACAAATAGACAGAGTGAAATAAAGGAAATATA	203311
QY	1134	TGGGCGCAAGAACTCTAAGGTTATTTGCAAAATCAATTGAAACCTTTAAAAAGAAAGAA	1193
Db	203310	TGGGCGCAAGAACTCTAAGGTTATTTGCAAAATCAATTGAAACCTTTAAAAAGAAAGAA	203251
QY	1194	TCACAGAGGCAATGAAAGACAGGGGAGAACAGGGGACAGAAACACTGTGGCCCAAGGA	1253
Db	203250	TCACAGAGGCAATGAAAGACAGGGGAGAACAGGGGACAGAAACACTGTGGCCCAAGGA	203191
QY	1254	GAACAAAAACAAGCTCTTAGACAGACAGAG- ---GAGAGAGAGAGAGTGAAGTGA	1309
Db	203190	GAACAAAAACAAGCTCTTAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGTGAAGTGA	203131
QY	1310	GACAGACAGAGAAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAG	1369
Db	203130	GACAGACAGAGAAAAAGACAGAGAGAGAGACAGAGACAGAGACAGAGAGGCGAGAG	203071
QY	1370	GGATTGAAAGAGAGAGAGGGGTGGAGAGACAGAGATATTGAGAGAGACTCAGAAAAA	1429
Db	203070	GGATTGAAAGAGAGAGAGGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAAA	203011
QY	1430	TAGCCGAGGGGAAACACAGAGAGATGAGAAAGACTCTGAGAAAAAACACAGAGCAAAAG	1489
Db	203010	TAGCCGAGGGGAAACACAGAGAGATGAGAAAGACTCTGAGAAAAAACACAGAGCAAAAG	202951
QY	1490	ATGGAAGAGAGATTCGAGGGGTGAACAGACGTGTGGAATGACCAAAATGCAAGAAAG	1549
Db	202950	ATGGAAGAGAGATTCGAGGGGTGAACAGACGTGTGGAATGACCAAAATGCAAGAAAG	202891
QY	1550	AAAGCAAGCAATCCAGGGCCCAAGAAATGTGACCCAGAGTTGTGTAAGAACCAAGATCCTT	1609
Db	202890	AAAGCAAGCAATCCAGGGCCCAAGAAATGTGACCCAGAGTTGTGTAAGAACCAAGATCCTT	202831
QY	1610	AAGGCTGGGGAGGCAAGGAGGGGCTGGCTTCCGAGACCCCTCCCATTTCTCC	1669
Db	202830	AAGGCTGGGGAGGCAAGGAGGGGCTGGCTTCCGAGACCCCTCCCATTTCTCC	202771
QY	1670	GGGCGAGGAGTTCAGAGTGAATTCGGACTGTGGTGGGGGTCTCTGGGGGTGAGAA	1729
Db	202770	GGGCGAGGAGTTCAGAGTGAATTCGGACTGTGGTGGGGGTCTCTGGGGGTGAGAA	202711
QY	1730	TAGGGGAGACAGAGAGACTTTGTCTAAGGCCCATTAGGCACCTATTGGCCGGAAATGT	1789
Db	202710	TAGGGGAGACAGAGAGACTTTGTCTAAGGCCCATTAGGCACCTATTGGCCGGAAATGT	202651
QY	1790	GGCCCAAGGAGCAGTGGGTGTATTAATCAAGCCCGGTGCCAGAGCCCAAGAGAGAGC	1849
Db	202650	GGCCCAAGGAGCAGTGGGTGTATTAATCAAGCCCGGTGCCAGAGCCCAAGAGAGAGC	202591
QY	1850	AGTGGCCAGAAAGCAAGGCTTAGAAAGTCTGCGGCTGAAGCTGTGGAGCAATGCCCAAC	1909
Db	202590	AGTGGCCAGAAAGCAAGGCTTAGAAAGTCTGCGGCTGAAGCTGTGGAGCAATGCCCAAC	202531
QY	1910	CCCTTACTGGGGGACAGGGCAATGAGACCTGTGTAGAGGTGGCTCAGACAGCAGGAAAG	1969
Db	202530	CCCTTACTGGGGGACAGGGCAATGAGACCTGTGTAGAGGTGGCTCAGACAGCAGGAAAG	202471

QY	1970	TAGAGAGTGTCTGTGGTCTCTGGACCCACATCTTTCTCTGTCCCTCTCTGCTGTGG	2029
Db	202470	GAGAGGTGTCTGTGGTCTCTGGACCCACATCTTTCTCTGTCCCTCTCTGCTGTGG	2024111
QY	2030	AGGCTGTAGATCTCCATCTTCTGAATTTCTATAGTCTCTGGGCTCTCAGCGCATGTCCGAT	2089
Db	202410	AGGCTGTAGATCTCCATCTTCTGAATTTCTATAGTCTCTGGGCTCTCAGCGCATGTCCGAT	2023511
QY	2090	GGTGGCCCGTCTTGTGGTTCTCTCTACCTGGGGAAATPAGGTAGGGAGGGAGAA	2149
Db	202350	GGTGGCCCGTCTTGTGGTTCTCTCTACCTGGGGAAATCAGGTAGGGAGGGAGGGAA	2022911
QY	2150	GTGGGTTAAGGCTCCCGGATTCGCTGTGGGCTCCCAACCTCTGTGACATTCGCATCCAG	2209
Db	202290	GTGGGTTAAGGCTCCCGGATTCGCTGTGGGCTCCCAACCTCTGTGACATTCGCATCCAG	2022311
QY	2210	GTGCAGCGCCCATGTGCTACAGCAAGACCCCTCGATGTGGTGTCTGTGCTCTGATCA	2269
Db	202230	GTGCAGCGCCCATGTGCTACAGCAAGACCCCTCGATGTGGGTCGTGTGCTCTGATCA	2021711
QY	2270	CAGCCTTGCTTCTGGGGGTCAAGGTAACTCTGGGGAAATPAGGTAGGGAGGTGTGGGATTT	2329
Db	202170	CAGCCTTGCTTCTGGGGGTCAAGGTAACTCTGGGGAAATPAGGTAGGGAGGTGTGGGATTT	2021111
QY	2330	GGGAGAGACTGTCTGTGGGCACTAGACGCTGTCCCTCTGGGGAACTGTGTAGGCTGGG	2389
Db	202110	GGGAGAGACTGTCTGTGGGCACTAGACGCTGTCCCTCTGGGGAACTGTGTAGGCTGGG	2020511
QY	2390	CATGACTCCGGGACCGGGGTGAATGTGATGTCTGTCTGTATCTTGTGGTGTGTGCATGTGA	2449
Db	202050	CATGACTCCGGGACCGGGGTGAATGTGATGTCTGTCTGTATCTTGTGGTGTGTGCATGTGA	2019911
QY	2450	TGTGGCCCTGTACTGTCCACGAGGTGTGTCTGGGGAGGGGGATATCCCTTTCACATATCAG	2509
Db	201990	TGTGGCCCTGTACTGTCCACGAGGTGTGTCTGGGGAGGGGGATATCCCTTTCACATATCAG	2019311
QY	2510	TGACTGTGGGAGGTGTGACATGACCCCTTGTAGGCTGTGTGTGTGGTTTGTGATTTGTGT	2569
Db	201930	TGACTGTGGGAGGTGTGACATGACCCCTTGTAGGCTGTGTGTGTGGTTTGTGATTTGTGT	2018711
QY	2570	GTGCAATTTAAGATGTGTGTGGCTCCACAGCTGTGTGGGTGAATGTGATPAGCATGGGG	2629
Db	201870	GTGCAATTTAAGATGTGTGTGGCTCCACAGCTGTGTGGGTGAATGTGATPAGCATGGGG	2018111
QY	2630	GTGTTCACGTGTGTGTGGCTGTGTGTGTGTGACTTGGCATTTGTATATGACTGCAG-----	2684
Db	201810	GTGTTCACGTGTGTGTGGCTGTGTGTGTGTGACTTGGCATTTGTATATGACTGCAGTCAT	2017511
QY	2685	-----GTATCTGCAG-----TTCTGTCTCTGAGGTCCCGGATTTG	2720
Db	201750	ATACATATCCCTGTATGTATATGACAGAGATACATATACATATCCCTGAGGTCCCGGATTTG	2016911
QY	2721	CGTGCACAACAAAGTGGTATCATCAACATGAAAGCTGTGACTGTGTGCTGTGACGCGAT	2780
Db	201690	CGTGCACAACAAAGTGGTATCATCAACGTTGAAAGCTGTGACTGTGTGCTGTGACGCGAT	2016311
QY	2781	TATGTGATTTGTGGCTGAGTGTGACGTTATATGATGCCGTATTTGTGACCGGTGTGACTACC	2840
Db	201630	TATGTGATTTGTGGCTGAGTGTGACGTTATATGATGCCGTATTTGTGACCGGTGTGACTACC	2015711
QY	2841	TGAACCTCTGTGTAGGGGTGACGTGATGTGACCTGTGTGTCTGTGTGTGAGGCCGTGTAA	2900
Db	201570	TGAACCTCTGTGTAGGGGTGACGTGATGTGACCTGTGTGTCTGTGTGTGAGGCCGTGTAA	2015111
QY	2901	TGCTACTGTATGTGTATGTGTGTCAGCTGTGTGTCTGTGAGATTTCTGTCTGTGCTGTAGGG	2960
Db	201510	TGCTACTGTATGTGTATGTGTGTCAGCTGTGTGTCTGTGAGATTTCTGTCTGTGCTGTAGGG	2014511
QY	2961	ATAGAGGGGTGCAGGGGTAGCTATCTCTGGAGATGTGGTGCAGGTGATCTGTGCAGTG	3020
Db	201450	ATAGAGGGGTGCAGGGGTAGCTATCTCTGGAGATGTGGTGCAGGTGATCTGTGCAGTG	2013911

QY	3021	TGTGCTGTGTGCACAAAGATATGTGGCAGTCTTGAACATCTGTGCACACACGGCAATCTGT	3080
Db	201390	TGTGCTGTGTGCACAAAGATATGTGGCAGTCTTGAACATCTGTGCACACACGGCAATCTGT	201331
QY	3081	GGCTGGCACTGAGCACTGTGGATAGAGGGTGTGCATCCCGCTTAGAGCTCCCGGGAGCGT	3140
Db	201330	GGCTGGCACTGAGCACTGTGGATAGAGGGTGTGCATCCCGCTTAGAGCTCCCGGGAGCGT	201271
QY	3141	GTGTACTGTGAGACAGAGCTGTATGTATGTAGCTGTGCACCTGTGAGGCAACATGGGCGTGTCT	3200
Db	201270	GTGTACTGTGAGACAGAGCTGTATGTATGTAGCTGTGCACCTGTGAGGCAACATGGGCGTGTCT	201211
QY	3201	GGAGAATCTGGCTGTGCTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	3260
Db	201210	GGAGAATCTGGCTGTGCTGTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	201151
QY	3261	TGCTAATGATATGTGTGTGTGCACAGGGCGATCAGCAAGAGGTAAAGAACAGGCGGGCGCGGTG	3320
Db	201150	TGCTAATGATATGTGTGTGTGCACAGGGCGATCAGCAAGAGGTAAAGAACAGGCGGGCGCGGTG	201091
QY	3321	GCTCAGCGCTGTATCCAGGCCCTTTTGGAGGGCCGAGGCGAGGCGATCACTGAGGTGCG	3380
Db	201090	GCTCAGCGCTGTATCCAGGCCCTTTTGGAGGGCCGAGGCGAGGCGATCACTGAGGTGCGG	201031
QY	3381	GAGATCGAGGCCAGGCTTGACCAACTGAGAACCCCGTCTTACTTAAATATCAAAAAT	3440
Db	201030	GAGATCGAGGCCAGGCTTGACCAACTGAGAACCCCGTCTTACTTAAATATCAAAAAT	200971
QY	3441	TAGCTGT	3500
Db	200970	TAGCTGT	200911
QY	3501	TGCTGTGAACCCGGAGGTGTGAGGTGTGGGTGTGAGCCGAGATGCGCCATTGCACTCCAGC	3560
Db	200910	TGCTGTGAACCCGGAGGTGTGAGGTGTGGGTGTGAGCCGAGATGCGCCATTGCACTCCAGC	200851
QY	3561	CTGGGCAACAAAGAGGAACTCCGCTCTGAAAGAAAAAGAAAAAGGGGTAAAGAA	3620
Db	200850	CTGGGCAACAAAGAGGAACTCCGCTCTGAAAGAAAAAGAAAAAGGGGTAAAGAA	200791
QY	3621	CCAGTGAATGGGCAAGGGAGGCACTGATATGAGGTGGGGCAATGATGTGTGTATGAGGTC	3680
Db	200790	CCAGTGAATGGGCAAGGGAGGCACTGATATGAGGTGGGGCAATGATGTGTGTATGAGGTC	200731
QY	3681	TGTGTGTGAGAGAGAGATTGACAGATTGAGAAAGGATGTTTTCACTGAGAAATTGAG	3740
Db	200730	TGTGTGTGAGAGAGAGATTGACAGATTGAGAAAGGATGTTTTCACTGAGAAATTGAG	200671
QY	3741	AAACTTAGGCTGTCTTCCCTCCATGTGTGCCCCCTTAAAGCTGAGCCCTTCTTCTGTGT	3800
Db	200670	AAACTTAGGCTGTCTTCCCTCCATGTGTGCCCCCTTAAAGCTGAGCCCTTCTTCTGTGT	200611
QY	3801	CCTGTGTTTGGAAACCTTAGCTCCGCGCCATGAGCTGTGACCCCACTCTTCTTCTCAACCA	3860
Db	200610	CCTGTGTTTGGAAACCTTAGCTCCGCGCCATGAGCTGTGACCCCACTCTTCTTCTCAACCA	200551
QY	3861	CGCCCTTAGGCAAGATCTAGTGTGACCCCGCTTAAGGCAACCCCTTGGGGCAGGCTC	3920
Db	200550	CGCCCTTAGGCAAGATCTAGTGTGACCCCGCTTAAGGCAACCCCTTGGGGCAGGCTC	200491
QY	3921	CACCCCTATTCTGTGGGTACTCTTGAAGACCCCTTCAAGATGAGAGC-TTTTTTTTTT	3979
Db	200490	CACCCCTATTCTGTGGGTACTCTTGAAGACCCCTTCAAGATGAGAGC-TTTTTTTTTT	200431
QY	3980	TTTTTTTTTGGAGACGTCTGTGCTCTCTCCAGGCTGTAGTGTGAGTGTGTGTGTGTGTGTGTGT	4039
Db	200430	TTTTTTTTTGGAGACGTCTGTGCTCTCTCCAGGCTGTAGTGTGAGTGTGTGTGTGTGTGTGTGT	200371
QY	4040	CTCACTGCAACCTGTGCGCTCCAGGTTCAAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4099
Db	200370	CTCACTGCAACCTGTGCGCTCCAGGTTCAAGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	200311
QY	4100	CTGGAGTTTACAGTGTGCGGCCAACACGCGTGGCTAATTTTGTGTCTTTAGTAGACAG	4159

Db 200310 CTGGAGTTAAGAGTGGCGGACCAACGCTGGCTAATTTTGTCTTTAGTATAGACAG 200251
 Qy 4160 GGTTCACCTTTTGGCCAGGCTGTCTCAAACTCCCACTCAGTGATCCGCCACT 4219
 Db 200250 GGTTTCACTTGTGGCCAGGCTGTCTCAAACTCCCACTCAGTGATCCGCCACT 200191
 Qy 4220 CGGCTCCCAAGAGTGGGGTTACAGCGGAGCACCGGCCAGCCCAAGTCAGAG 4279
 Db 200190 CGGCTCCCAAGAGTGGGGTTACAGCGGAGCACCGGCCAGCCCAAGTCAGAG 200131
 Qy 4280 CTCTTTAAGAGACTCTAACATGTAACTCTGAGCCCTGAGCCCTAAGTCAATTC 4339
 Db 200130 CTCTTTAAGAGACTCTAACATGTAACTCTGAGCCCTGAGCCCTAAGTCAATTC 200071
 Qy 4340 ACCCTTCCTGCTCCAGGCTGACCCCACTCACTGAGGCTGACCCCACTTTTGA 4399
 Db 200070 ACCCTTCCTGCTCCAGGCTGACCCCACTCACTGAGGCTGACCCCACTTTTGA 200011
 Qy 4400 CAGTTCCATCCCTAAAGCCCTGCTCCCTCCCTCCAGGCTCCAGCCCAAGCTT 4459
 Db 200010 CAGTTCCATCCCTAAAGCCCTGCTCCCTCCCTCCAGGCTCCAGCCCAAGCTT 199951
 Qy 4460 TGGCACTACCCCTGAGCTTGTCCAGAACTCTGTAACCAATTTTACCCTCACTAGT 4519
 Db 199950 TGGCACTACCCCTGAGCTTGTCCAGAACTCTGTAACCAATTTTACCCTCACTAGT 199891
 Qy 4520 CTACCAATTCAGAAATCTGTGAGGTCAGTTAGATGACGTACGTAACCTTACGCTG 4579
 Db 199890 CTACCAATTCAGAAATCTGTGAGGTCAGTTAGATGACGTACGTAACCTTACGCTG 199831
 Qy 4580 GGGCTCTCTCTGAGCTTGAAGCCGCTGAGAGGTCACCTTATTTCCAGGCTT 4639
 Db 199830 GGGCTCTCTCTGAGCTTGAAGCCGCTGAGAGGTCACCTTATTTCCAGGCTT 199771
 Qy 4640 GCCCTGCTCCCTGAGATGTCAAGACCACTCTAGCTGTGCTGCTCTTGAAGCT 4699
 Db 199770 GCCCTGCTCCCTGAGATGTCAAGACCACTCTAGCTGTGCTGCTCTTGAAGCT 199711
 Qy 4700 GAAACCAACCCCAAGCCCAAGCCGCTGAGCCCGGCCCAACCAATTTTCCGTTCC 4759
 Db 199710 GAAACCAACCCCAAGCCCAAGCCGCTGAGCCCGGCCCAACCAATTTTCCGTTCC 199651
 Qy 4760 AGAGCATGTTCTGCGCAACATGATGTTCTGTGACCAACCCCTTAAACCCGCTC 4819
 Db 199650 AGAGCATGTTCTGCGCAACATGATGTTCTGTGACCAACCCCTTAAACCCGCTC 199591
 Qy 4820 TGGAGAGCAACGAGACCTGGAGCTGGGGCCGGGGAAGACGCGGTCGATGACAGAG 4879
 Db 199590 TGGAGAGCAACGAGACCTGGAGCTGGGGCCGGGGAAGACGCGGTCGATGACAGAG 199531
 Qy 4880 CAGCCGATCATCATGATGATCGAATCGATATGACACCAAGCCGTTGGAGCGCGCT 4939
 Db 199530 CAGCCGATCATCATGATGATCGAATCGATATGACACCAAGCCGTTGGAGCGCGCT 199471
 Qy 4940 GTTGCTAAGCGCCCAACGAGCTTCTGCGGGCGGTGTGTGATGATCAACAGTGGCT 4999
 Db 199470 GTTGCTAAGCGCCCAACGAGCTTCTGCGGGCGGTGTGTGATGATCAACAGTGGCT 199411
 Qy 5000 C 5000
 Db 199410 C 199410

RESULT 8
 AC130782/c
 LOCUS
 DEFINITION Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
 AC130782 200792 bp DNA linear HTG 22-NOV-2002
 ordered pieces.
 AC130782
 VERSION AC130782.2 GI:25167101
 HTG; HTGS PHASE2; HTGS DRAFT.
 KEYWORDS
 SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
 REFERENCE
 1 (bases 1 to 200792)
 AUTHORS
 Banister, N., Antonellis, A., Ayala, K., Beckstrom-Sternberg, S. M.,
 Benjamin, B., Blakeley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S.,
 Carls, K., Coleman, B., Engle, J., Grant, S., Guan, X., Gupta, J.,
 Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E.,
 Latic, P., Lee, L.-S., O., Legaspi, R., Maduro, Q. L., Maduro, V. B.,
 Margulies, E. H., Maestri, C., Maestri, B., McDowell, J.,
 Pasquigian, C., Pearson, R., Portnoy, M. E., Prasad, A.,
 Reddix-Dugue, N., Schandier, K., Scheller, M. G., Sison, C.,
 Stanciu, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L.,
 Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.
 NISC Comparative Sequencing Initiative
 UNPUBLISHED
 2 (bases 1 to 200792)
 REFERENCE
 Green, E. D.
 TITLE
 Direct Submission
 Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 200792)
 REFERENCE
 Green, E. D.
 TITLE
 Direct Submission
 Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Nov 22, 2002 this sequence version replaced gi:22218452.
 COMMENT
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc.zoe@nigr1.nih.gov
 ----- Project Information
 Center project name: dmz
 Center clone name: 355A20

The sequence data in this record represents an 'enhanced'
 version of a Phase 2 submission. Specifically, the indicated
 order and orientation of each sequence contig has been
 established using one or more of the following: read-pair
 data from individual subclones, overlaps with neighboring
 clones, alignment with available reference sequence (e.g.,
 human), and/or confirmation by PCR testing. In addition,
 the sequence assembly is based on at least 8X average
 coverage in Q20 bases and has been reviewed to rule out
 gross misassemblies, the low-quality ends of sequence
 contigs have been trimmed away, and each base is associated
 with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 19656 bases at least Q40
 Consensus quality: 19783 bases at least Q30
 Consensus quality: 19887 bases at least Q20
 Insert size: 21500; agarose-fp
 Insert size: 19692; sum-of-contigs
 Quality coverage: 9.03x in Q20 bases; agarose-fp
 Quality coverage: 9.73x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 28306: contig of 28306 bp in length
 * 28307 28406: gap of unknown length
 * 28407 37856: contig of 9450 bp in length
 * 37857 37956: gap of unknown length

*	37357	73522:	cont'g of 35566 bp in length
*	73523	73622:	gap of unknown length
*	73524	83567:	cont'g of 9945 bp in length
*	83568	83667:	gap of unknown length
*	83569	88811:	cont'g of 5150 bp in length
*	88810	88917:	gap of unknown length
*	88918	125511:	cont'g of 36694 bp in length
*	125512	125711:	gap of unknown length
*	125712	158879:	cont'g of 34168 bp in length
*	158880	159797:	gap of unknown length
*	159880	174698:	cont'g of 14719 bp in length
*	174699	174798:	gap of unknown length
*	174799	186382:	cont'g of 11584 bp in length
*	186383	186482:	gap of unknown length
*	186483	193444:	cont'g of 6662 bp in length
*	193445	193444:	gap of unknown length
*	193446	199365:	cont'g of 5919 bp in length
*	199366	199465:	gap of unknown length
*	199466	200792:	cont'g of 1329 bp in length

FEATURES	Location/Qualifiers
source	1. .200792

```

misc_feature
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /clone="CH251-355A20"
    /clone_1b="CH251"
    1..28306
    /note="assembly_fragment"
    clone_end:sp6
    vector_side:left"

misc_feature
    28407..37856
    /note="assembly_fragment"
    37957..73522
    /note="assembly_fragment"
    73623..83567
    /note="assembly_fragment"
    83668..88817
    /note="assembly_fragment"
    88918..125611
    /note="assembly_fragment"
    125712..159879
    /note="assembly_fragment"
    159980..174698
    /note="assembly_fragment"
    174799..186382
    /note="assembly_fragment"
    186483..193344
    /note="assembly_fragment"
    193445..199363
    /note="assembly_fragment"
    199464..200792
    /note="assembly_fragment"
    clone_end:T7
    vector_side:right"

```

ORIGIN

Query Match	Score	DB 2	Length
82.1%	4106.4	200792	

Best Local Similarity 97.3%; Pred. No. 0;
Matches 4268; Conservative 0; Mismatches 66; Indels 52; Gaps 7;

QY	654	GGCCATGATCATCTTTCTTGACATATGCTGATGTGACAAAGTACTCTAAAGCCATCAGACTCTTA	713
Db	78497	GTCTATGATCATCTTTCTTGACATATGCTGATGTGACAAAGTACTCTAAAGCCATCAGACTCTTA	78438

QY	714	CCCTTTAATATACAGTTTGGGCGACGACCGTGGCTCATGCTGTAATTCAGACCTTT	773
Db	78437	CCCTTTAATATACAGTTTGGGCGACGACCTGGCTCATGCTAATTCAGACCTTT	78378

QY 774 GGGAGCGCAGAGTGGCTGATATCATCTTGAAGGCCAGAGTTTGAAGACCAAGCCTTGGCCACAT 833

Db 78377 GGGAGCGCAGAGTGGCTGATATCATCTTGAAGGCCAGAGTTTGAAGACCAAGCCTTGGCCACAT 78318

834 GGTGAACTCTGCTTACTAAAAAATCAGCCGGTCTCGT 893

Db 78317 GGTAAGCTCTGTCTTACT-----AAAAAAAAAAAAAAAAATCAGCCGGGTGTCCG 78264

894 GGGGCACACCTGTAAATCCACGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAACCT 953

Db 78263 GGGCACACCTGTATCCAGCTACGCCGAGGCTGAGGCACGAGAGTCACTTGAACCT 78204

954 GGAGCGGAGGTTGCAGTGGGCCGAGTCAATCACCGCCCTCCAGCCTGGGCGACAGAG 1013

DD /0203 0606G1.06LAD1.0606C0A0H1.CA0C1.CA0C0C0C1.C0C0C0C1.0606C0A0H0GND /0217

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

CCNA 1074 *****CTCNCACATBAGGAGTGAATTA 1131

Db 78083 TATACCTAAAAAATGCTGTCAACCAATAGAGCAGAACTGAATTAAGGAAATA 78024

QY 1132 AATGGCCAAGACTCTAAGTATATTGACCAATCATTCAGAACCTTAAAAAAGAAG 1191

Db 78023 AATGGGCCAAGACTCTAAGTATATTGACCAATGATTCAGAACCTTTAAAAAAGAAAG 77964

1192 AATCAGAGGCATAGAAAGCAGGGAGGAACAGGAGACAGAAACA CCGTGTGCCCAAG 1251

DB / 1963 AATCAGAGGCAIAGAAAGACAGGAGAACAGGAGACAGAAACACCCGAGCCCAAG / 1904

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

1308 1347

Db 77843 GAGACAGACAGAAAAGACAGAGAGAGAGACAGACAGACGCGAGAGGATAGAAAG 77784

QY 1348 -----CAGAGACAGAGAGCGGAGGATAGAAAGAGAGAGAGGGTGA 1394

Db 77783 AGAGAGAGCGGTGAGAGAGACAGAGAGCGCGAGAGCGATAGAAAGAGAGAGAGCGGTGA 77724

QY 1395 GAGAGACGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGAGAA CCA CAGAGAGA 1454

Db 77723 GAGAGACACGAGATATGAGAGAGACTCAGAAAGATAGCCGAGGAGAACCAAGAGAGA / 1664

27	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	

[illegible]

77503 ACAGACAGTGTGATGAGCAAAATGCAGAGGAAGCAAGCAATCCAGGCGCCAAGA 77544

1575 ATAGTGACCCAGAGTTGGTGAGAACCCAGATCCTTAAGCTGGGGAGGCGAGGAGGGG 1634

Db 77543 ATAGTGACCCAGAGTTGGTGAGAACCCAGATCCTTAAGGCTGGGGAGCCAGGGAAGGGG 77484

QY 1635 CTGGCCTGGCTTCCGGAGACCCCTCCCATTTCTCCGGCCAGGGAGGTAGGAGTGACAT 1694

D_b 77483 CTGGCCTGGCTCCAGAGACCCCTCCCATTTCTCCCCGCCAGGAGGCAGGAGTGACAT 77424

1695 1C CCGAC1GGG1GGGGGG1G1C1C1GGGG1GGAGGA1AGGGGGAGG1CAAGAGGAGC1A11G1C 1/54

[illegible]

Dh 77363 TAAAGGATACCTCATTTGGGAGTGTGAT 77304

1815 AACTCAGGCCCGGTGCCCAGAGCCAGGAGGAGGCACTGGCCAGGAAGGCACAGGCCCTGA 1874

Db 77303 AACTCAGGCCCGGTGCCCAGAGCCAGAGGAGGCACTGGCCAGGAAGGCACAGGCTGA 77244

1875 G A A G T C T G C G G C T G A G C T G G A G C A A T C C C C C A C C C C T A C T G G G G A C A G G C A A G T 1934

Db 77243 GAAGCTGCGGCTGAGCTGGAGCAAAATCCCCACCCCTACTGCGGGAGACAGGGCAAGT 77184
QY 1935 GAGACCTGGTGAAGGTGGCTACAGACGACAGGAAAGAGAGGTGTCTGTGGCTCTTGCAAC 1994
Db 77183 GAGACCTGGTGAAGGTGGCTACAGACGACAGGAAAGAGAGGTGTCTGTGGCTCTTGCAAC 77124
QY 1995 CACATCTTTCTCTGTCCCTCTCTTGCCCTGTGTGGAGGCTGTGAACCTCATCTTCTGA 2054
Db 77123 CACATCTTTCTCTGTCCCTCTCTTGCCCTGTGTGGAGGCTGTGAACCTCATCTTCTGA 77064
QY 2055 ATTCTAATAGTCCGTGGTCTCAGCGCAGTCCGATGTGCGCCGTCTGTGTCTCTCT 2114
Db 77063 ATTCTAATAGTCCGTGGTCTCAGCGCAGTCCGATGTGCGCCGTCTGTGTCTCTCT 77004
QY 2115 CTACCTGGGGAAATPAGGTAGGGAGGGAGGGAAATGGGTTAAAGGCTCCCGGATCC 2174
Db 77003 CTACCTGGGGAAATPAGGTAGGGAGGGAGGGAAATGGGTTAAAGGCTCCCGGATCC 76944
QY 2175 CTGGGCGTCCCAACCTCTGACATTTCCCATCCAGGTGACAGCGGCATGGCTACAGAG 2234
Db 76943 CTGGGCGTCCCAACCTCTGACATTTCCCATCCAGGTGACAGCGGCATGGCTACAGAG 76884
QY 2235 ACCCCCTGAGTGGGTGTCTGTGTCTGTGATCAGCCTTGCTTGAGGGGTCAAG 2294
Db 76883 ACCCCCTGAGTGGGTGTCTGTGTCTGTGATCAGCCTTGCTTGAGGGGTCAAG 76824
QY 2295 TAACCAAGACTCTGGGGTGGAGGGGTGGGATTTGGAGGACTGTCTCTGCGGACTAG 2354
Db 76823 TAACCAAGACTCTGGGGTGGAGGGGTGGGATTTGGAGGACTGTCTCTGCGGACTAG 76764
QY 2355 AGGCGCTGTCCCTGGGGGAACTGTGTGAGCCTGGGGCACTGCGGGACCGGGGAAATGT 2414
Db 76763 AGGCGCTGTCCCTGGGGGAACTGTGTGAGCCTGGGGCACTGCGGGACCGGGGAAATGT 76704
QY 2415 GAGTCTGTCTGACTTGT 2474
Db 76703 GAGTCTGTCTGACTTGT 76644
QY 2475 GTGTGCGGAGAGGGGATGCTTTTCCATATCAGGTGACTGTGCGGACAGGTGCACTGAC 2534
Db 76643 GTGTGCGGAGAGGGGATGCTTTTCCATATCAGGTGACTGTGCGGACAGGTGCACTGAC 76584
QY 2535 CCTTTGAGGCTGT 2594
Db 76583 CCTTTGAGGCTGT 76524
QY 2595 CACAGCTGT 2654
Db 76523 CACAGCTGT 76464
QY 2655 GTGTGACTGT 2714
Db 76463 GTGTGACTGT 76404
QY 2715 GGAATGT 2774
Db 76403 GGAATGT 76344
QY 2775 GGCATTAATGT 2834
Db 76343 GGCATTAATGT 76284
QY 2835 ACTACCTGAAGCTGT 2894
Db 76283 ACTACCTGAAGCTGT 76224
QY 2895 TGTAAATGT 2954
Db 76223 TGTAAATGT 76166
QY 2955 GAGGAGATAGAGGTGACAGGGGTAGCTATCTGTGGAGATGGGTGCGAGGTGACTGACTT 3014
Db 76165 GAGGAGATAGAGGTGACAGGGGTAGCTATCTGTGGAGATGGGTGCGAGGTGACTGACTT 76106

QY 3015 GAGT 3074
Db 76105 GAGGAGT 76046
QY 3075 ATCTGT 3134
Db 76045 ATCTGT 75986
QY 3135 GAGGT 3194
Db 75985 GAGGT 75926
QY 3195 GT 3254
Db 75925 GT 75866
QY 3255 GTGAGTTCTGTGAATGT 3314
Db 75865 GTGAGTTCTGTGAATGT 75806
QY 3315 GCGGT 3374
Db 75805 GCGGT 75746
QY 3375 GT 3434
Db 75745 GT 75686
QY 3435 AAAAATTAAGCTGT 3494
Db 75685 AAAAATTAAGCTGT 75626
QY 3495 GAAAAATGT 3554
Db 75625 GAAAAATGT 75566
QY 3555 TCCAGCTGT 3614
Db 75565 TCCAGCTGT 75506
QY 3615 TAAGAACCTGT 3674
Db 75505 TAAGAACCTGT 75446
QY 3675 TAGGT 3734
Db 75445 TAGGT 75386
QY 3735 ATTCAAGAACTGT 3794
Db 75385 ATTCAAGAACTGT 75326
QY 3795 CTTGT 3854
Db 75325 CTTGT 75266
QY 3855 CAAACAGCCCTTAAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3914
Db 75265 CAAACAGCCCTTAAGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 75206
QY 3915 AGGCTCAACCCCTTAATGT 3974
Db 75205 AGGCTCAACCCCTTAATGT 75149
QY 3975 TTTTGT 4034
Db 75148 TTTTGT 75089
QY 4035 CTCGGCTCACTGT 4094
Db 75088 CTCGGCTCACTGT 75029

QY 4095 AGTAGCTGGATTACAGTGTGCGGCAACACGCTGGCTAATTTTGTGTCTTTAGTAGA 4154
 Db 75028 AGTAGCTGGATTACAGTGTGCGGCAACACGCTGGCTAATTTTGTGTCTTTAGTAGA 74969
 QY 4155 GACAGGGTTTACCTGTTTGGCCAGAGGTGTCTCAAACTCCAACTCAGGTATCCGCC 4214
 Db 74968 GACAGGGTTTACCTGTTTGGCCAGAGGTGTCTCAAACTCCAACTCAGGTATCCGCC 74909
 QY 4215 CACCTGGGCTCCAGAGTGTGGGTTTACAGGCGTAGGCAACGCGCCAGCCCAAGT 4274
 Db 74908 CACCTGGGCTCCAGAGTGTGGGTTTACAGGCGTAGGCAACGCGCCAGCCCAAGT 74849
 QY 4275 CAGAGCTTTTATAGAGACTTACATGTATACCTGACCCCTGACCTTAAGTCAAT 4334
 Db 74848 CAGAGCTC--TATAGAGACTTACATGTATACCTGACCCCTGACCTTAAGTCAAT 74791
 QY 4335 TCCAAACCCCTTCTGCTCTCAAGCCTGACCCCACTCACTGAGGCTGACCCCACTTCTT 4394
 Db 74790 TCCAAACCCCTTCTGCTCTCAAGCCTGACCCCACTCACTGAGGCTGACCCCACTTCTT 74731
 QY 4395 GAGACAGTTTCATCCCTTAAAGCCCTGATCCCTCCATCCCGAGGCTCAGCCGCCAC 4454
 Db 74730 GAGACAGTTTCATCCCTTAAAGCCCTGATCCCTCCATCCCGAGGCTCAGCCGCCAC 74671
 QY 4455 AGCTTTGGCACTACCCCTGAGCTTGTCCAGAAATCCCTGTAACCAATTTTACCTTCATG 4514
 Db 74670 AGCTTTGGCACTACCCCTGAGCTTGTCCAGAAATCCCTGTAACCAATTTTACCTTCATG 74611
 QY 4515 TAGTTCTAGCAATTCAGGAATCTGTGAGGTCAAGTTAGATCCAGTAACCTTACCTGA 4574
 Db 74610 TAGTTCTAGCAATTCAGGAATCTGTGAGGTCAAGTTAGATCCAGTAACCTTACCTGA 74551
 QY 4575 GCCCTGGGCTGTGCTTGAAGCTTGAAGCTGGGCTGAGAGTGCACCTTATTTCCAG 4634
 Db 74550 GCCCTGGGCTGTGCTTGAAGCTTGAAGCTGGGCTGAGAGTGCACCTTATTTCCAG 74491
 QY 4635 GCCCTGGGCTGTGCTTGAAGCTTGAAGCTGGGCTGAGAGTGCACCTTATTTCCAG 4694
 Db 74490 GCCCTGGGCTGTGCTTGAAGCTTGAAGCTGGGCTGAGAGTGCACCTTATTTCCAG 74431
 QY 4695 AGTCTGAATCCCAACCCCGAGCCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 4754
 Db 74430 AGTCTGAATCCCAACCCCGAGCCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 74371
 QY 4755 TTCCAGAGCATTTTCGCGCAATGATGTTTCCGTGACACACCCCTCTAACAACGCTG 4814
 Db 74370 TTCCAGAGCATTTTCGCGCAATGATGTTTCCGTGACACACCCCTCTAACAACGCTG 74311
 QY 4815 CCTCTGGAGCAACGAGACTTGGAGCTGGGCGCGGAGAGACGCGCGGTGAGTAC 4874
 Db 74310 CCTCTGGAGCAACGAGACTTGGAGCTGGGCGCGGAGAGACGCGCGGTGAGTAC 74251
 QY 4875 AGCAGAGCCGCTCATCATGATCCGACTGCGATATGACACCCAGCCGTGGAGGCC 4934
 Db 74250 AGCAGAGCCGCTCATCATGATCCGACTGCGATATGACACCCAGCCGTGGAGGCC 74191
 QY 4935 GCGCTGTGTAGAGCCCAACGAGCTCTACTGGGCGCGGCTGTTGGTGCATCCAGCTG 4994
 Db 74190 GCGCTGTGTAGAGCCCAACGAGCTCTACTGGGCGCGGCTGTTGGTGCATCCAGCTG 74131
 QY 4995 CTGCTC 5000
 Db 74130 CTGCTC 74125

RESULT 9
 AC130188/c 176647 bp DNA linear HTG 14-NOV-2002
 LOCUS Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered
 DEFINITION pieces.
 ACCESSION AC130188
 VERSION AC130188.2 GI:24960890
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.

SOURCE
 ORGANISM
 Papio anubis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Papio.
 REFERENCE
 AUTHORS
 1 (bases 1 to 176647)
 Akher, N., Antonellis, A., Ayele, K., Beckett, S., Sternberg, S. M.,
 Benjamin, B., Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S.,
 Carriaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J.,
 Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E.,
 Latic, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B.,
 Margulies, E. H., Masello, C., Maskeri, B., McDowell, J.,
 Padgugirigan, C., Pearson, R., Portnoy, M. E., Prasad, A.,
 Reddi-Dugue, N., Schandler, K., Schueler, M. G., Sison, C.,
 Stancir, P., S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L.,
 Werber, K. D., Wiggins, L., Young, A. and Green, E. D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 176647)
 Green, E. D.
 Direct Submission
 Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 176647)
 Green, E. D.
 Direct Submission
 Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Nov 14, 2002 this sequence version replaced gi:22138439.
 COMMENT
 Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@phri.nih.gov
 Project Information
 Center project name: dar
 Center clone name: 421P03

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 171127 bases at least Q40
 Consensus quality: 173424 bases at least Q30
 Consensus quality: 174803 bases at least Q20
 Insert size: 18800; agarose-fp
 Insert size: 17547; sum-of-contigs
 Quality coverage: 8.76x in Q20 bases; sum-of-contigs
 Quality coverage: 9.39x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 6526: contig of 6526 bp in length
 * 6527 6626: gap of unknown length

*	6627	18812	contig of 12186 bp in length
*	18813	18812	gap of unknown length
*	18913	61780	contig of 42668 bp in length
*	61781	61880	gap of unknown length
*	97068	97068	contig of 35188 bp in length
*	97168	97168	gap of unknown length
*	97169	125206	contig of 28038 bp in length
*	125207	125306	gap of unknown length
*	125307	127021	contig of 1715 bp in length
*	127022	127121	gap of unknown length
*	127122	129387	contig of 2266 bp in length
*	129388	129487	gap of unknown length
*	129488	132102	contig of 2815 bp in length
*	132403	132402	gap of unknown length
*	132403	135735	contig of 3333 bp in length
*	135736	135835	gap of unknown length
*	135836	143558	contig of 7723 bp in length
*	143559	143658	gap of unknown length
*	143659	170182	contig of 26524 bp in length
*	170183	170282	gap of unknown length
*	170283	176647	contig of 6365 bp in length

FEATURES

```

/misc_featured
1. .6526
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-421P3"
/clone_1fb="RP41"
1. .6526
/note="assembly_fragment
clone_end:SP6
vector_side:left"
6627. .18812
/note="assembly_fragment"
18913. .61780
/note="assembly_fragment"
61881. .97068
/note="assembly_fragment"
97169. .125706
/note="assembly_fragment"
116911. .176647
/note="clone overlaps with Genbank Accession Number
AC130786 clone RP41-94H15 (center project name ds8)"
125307. .127021
/note="assembly_fragment"
127122. .125387
/note="assembly_fragment"
129488. .133303
/note="assembly_fragment"
132403. .135735
/note="assembly_fragment"
135836. .143558
/note="assembly_fragment"
143659. .170182
/note="assembly_fragment"
170283. .176647
/note="assembly_fragment
clone_end:17
vector_side:right"

```

ORIGIN

Query Match	20.7%	Score 1035.2;	DB 2;	Length 176647;
Best Local Similarity	88.1%;	Pred. No. 1.3e-260;		
Matches 1256;	Conservative 0;	Mismatches 123;	Indels 47;	Gaps 10;

Qy	Db	Qy	Db
658	1410	718	1350
ATGATCACTCTTGACATATCTATATGACAAGAACCTAAAGCATACAGACTTAACCT	ATCATCCCTTGATTATGTATGTGTGTGACAGATACCTAAAGCATATGACACTCCACCT	TTAAATATGACGTTTGCGGCACGACCGCTGCTCATGCTGTAATTCACGACATTTGGGA	TTAAATATGACGTTTGCGGCACGACCGCTGCTCATGCTGTAATTCACGACATTTGGGA

Qy	778	GGCAGAGTGGGTAAATACCTTGAGGCGAGGAGTTTGAACCACTCGGCCCAATGGTG	837
Db	1290	GCGAGAGTGGGTAAATACCTTGAGGCGAGGAGTTTGAACCACTCGGCCCAATGGTG	1231
Qy	838	AAACTCGTCTTTTACTAAAAAATCAAGCCGGGTGTCTGGGG	897
Db	1230	AAATCCCATCTTACT-----AAAAAATCAAGCCGGGCATGTGGGG	1176
Qy	898	CACACTCTTAATCCCACTATGCTGAGGCTGAGGCGAAGAGTCACTTGA--ACCTGG	955
Db	1175	--CACTGTATATCCCACTACTCCAGAGGCTGAGGCGTGAAGAGTCCGTTGAACACAGG	1118
Qy	956	AGGCGGAGGTTTGCAGTGGGCGAGATCAATCAACCCGCCCTCCAGCCGGGCGAAGAGCA	1015
Db	1117	AGGAGAGGTTTGCAGTGGGCGGAGATCAACCACTGCCCTCCAGCCGGGCGAAGAGCA	1058
Qy	1016	AGACTCTGTCTCAAAATTAATTAATAACAAACGAACAGGATTTGTTGTACTTAGTTA	1075
Db	1057	AGACTCTGTCTCAAAATTAATTAATAACA----AACAAACAGTTTGTATTATCTTAGGTA	1002
Qy	1076	TATCT-----AAAAAATATCTGTCAACAAATAGAGCAGAAGTGAATTAAGAAAAATA	1132
Db	1001	TACTTAATAAAAAAATATCTGTCAACAGATAGAGCAGAAGTGAATTAAGAAAAATA	942
Qy	1133	ATGGGCCCAAGACTCTTAAGGTAATTTTGGACAAATCATTCGAACCTTTAAAAAAGAA	1192
Db	941	ACAGGCTGAGAACTTAAGGTATATTGGACAAATCATCTCGAATCTTTAAAAAAGAGAA	882
Qy	1193	ATCACAGAGGCAATGAAGAAC---AGGAGGAACAGGGAGCAGAAACACTGTGGCCC	1248
Db	881	ATCACAGAGGCAATGAAGAACAGAGACGAGAAATGGGGAGACAGCACACTGTGGCCC	822
Qy	1249	AAGGAGAACAAACCAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGTGAAGTGA	1308
Db	821	AAGGAGAGAGAAACAAGGCTCTTAAGCA-----GGAAGAGAGAGAGAGAGAGAG	767
Qy	1309	AGACAGCACAGAAAAAGACAGAGAGAGAGACAGACAGAGAGCACAGAGGCCAGAA	1368
Db	766	CGCACTTGACACAGAGAGCACGCCAGACAGAGACAGAGAGACAGAGAGTGA	707
Qy	1369	GGGATAGAAAAAGAGA---GAGGGGTGAGAGAGACAGAGATTTTGAAGAGACTGAGA	1425
Db	706	GGGAGAGAGAAAGAGAGGGGATGAGAGAGACAGAGATTTTGAAGAGACTGAAA	647
Qy	1426	AAAGTACCCGAGGGAGAACACAGAGAGATGGAAGAACTCTG-----	1469
Db	646	AAAGCAGCCCTGGGAGAACACAGAGGAGATGGAAGAACTCTGAAAAAAGAAAA	587
Qy	1470	-AGAAAAACCAAGACCAAGATGGAAGAGAGATGAGAGGTGAACAGACTGTGG	1528
Db	586	AAAGAAAAAACAGAGCAAAAGATGGAAGAGAGATGAGAGCGGGAACAAGTGTGTGG	527
Qy	1529	AATCAGCAAAATCAGAGAGAAAGCAAGCAATCCAGGCCCAAGAAATATGTACCCAGAG	1588
Db	526	AATCAGCAAAACGAGAGAAAGAAAGCAAGCAATCCAGGACACAGAAACAGTGAACAGAG	467
Qy	1589	TTGGTGAAGAACAGATCTTTAAGGCTGTGGGAGGCGAGGAAAGGGCTGTGGCTTCC	1648
Db	466	TTGGTGAAGAACAGATCTCTTAAGGCTGTGGGAGGCGAGGAAAGGGCTGTGTGCTTCC	407
Qy	1649	GGAGACCCCTCCCATTTCTCGGGCCAGGGAGGTAGGAGATGACATTCGGACTGGGTGG	1708
Db	406	GGAGACCCCTCCCATCTCTCCGAGCCAGGAGGTAGGAAATGTGTAATTCGGGAACAGGTGG	347
Qy	1709	GGGTTGCTCTGGGGGTGAGATAGGGGGAGCAGAGAGAGCTATTGCTTAAGCCCGATAGG	1768
Db	346	GGGTTGCTCTGGGGGTGAGATAGGGGGAGCAGAGAGAGGATTTGCTTAAGGCCCAATAGG	287
Qy	1769	CACCTCAATTCACCGGAGATGTGGCCCAAGGAGCAGTGGGTGTTATTAATCAGAGCCGGT	1828
Db	286	CACCTCAATTCACCGGAGATGTGCCCCAGGAGCAGTGGGTGTTATTAATCAGAGCCGGT	227
Qy	1829	GCCCAAGGCCCAGAGAGGCAATGGCCAGAGAAAGCAACAGGCTGAGAAAGTCTTGGCTG	1888

|||||
Db 226 GCCCAGAGCCGAGAGGCGGTGGCCAGAGGTGCAGCCCTGAGAAATCCGTGGCTG 167
1889 AGCTGGAGAGCAATCCGCCACCCCTTACCTGGGGGACAGGGCAGTGAAGCTGGAGG 1948
Db 166 AACTGGAAGGAATCCGCCACCCCTTACCTGGGGGACAGGGCAGTGAAGCTGGAGG 107
Qy 1949 GTGGCTCAGCAGCAGGAGAGAGGTGTCTGTGCGTCTGACCCACATCTTTCTCTG 2008
Db 106 GTGGCTCAGCAGCAGGAGAGAGGTGTCTGTGCGTCTGACCCACATCTTTCTCTG 47
Qy 2009 TC--CCCTCTTCCCTGTCTGTGAGCGTGTGACCTCTTATCTTCT 2052
Db 46 TCTTCCCTCTCCGCCCGTGTGAGGTGTCTGACTCTTATCTTCT 1
RESULT 10
AC137591/c 32799 bp DNA linear PRI 05-DEC-2002
LOCUS Homo sapiens chromosome X clone XX-B6cos map Xp22-PAR, complete
DEFINITION sequence.
AC137591
VERSION AC137591.1 GI:25446705
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Rump,A., Rosenthal,A., Drescher,B. and Schatkevov,R.
TITLE 1 (bases 1 to 32799)
JOURNAL Submitted (07-MAY-1997) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE
AUTHORS Lagemann,D. and Platzter,M.
TITLE 2 (bases 1 to 32799)
JOURNAL Submitted (26-NOV-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
REFERENCE
AUTHORS Lagemann,D. and Platzter,M.
TITLE 3 (bases 1 to 32799)
JOURNAL Submitted (05-DEC-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany
COMMENT

Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de

Project Information
Center project name: B6
Center clone name: XX-540F24

Summary Statistics
Sequencing vector: pUC18, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 32765 bases at least Q40
Consensus quality: 32770 bases at least Q30
Consensus quality: 32770 bases at least Q20
Quality coverage:13.12x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source
Location/Qualifiers
1..32799
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp22-PAR"
/clone="XX-B6cos"
537..642
/note="single stranded/single chemistry region"
2807..2820
/note="single stranded/single chemistry region"
2913..2927
/note="low quality region"
3166..3173
/note="single stranded/single chemistry region"
5131..5234
/note="single stranded/single chemistry region"
5235..5317
/note="single clone coverage"
5266..5317
/note="low quality region"
5266..5317
/note="single clone coverage"
5366..5431
/note="single clone coverage"
5432..5518
/note="single stranded/single chemistry region"
6151..6395
/note="single stranded/single chemistry region"
6303
/note="single stranded/single chemistry region"
6382
/note="single stranded/single chemistry region"
7106..7137
/note="single stranded/single chemistry region"
9265..9528
/note="single stranded/single chemistry region"
9669
/note="T substituted in clone: XX-P99cos"
10317
/replace="c"
/note="deleted in clone: XX-540F24"
10897
/replace=""
/note="G substituted in clone: XX-1Bcos"
11087..11090
/replace="t"
/note="deleted in clone: XX-540F24"
11259
/replace=""
/note="T substituted in clone: XX-540F24"
11892..11900
/replace="c"
/note="deleted in clone: XX-1Bcos"
12197..12200
/replace=""
/note="deleted in clone: XX-1Bcos"
12256..12259
/replace=""
/note="deleted in clone: XX-1Bcos"
12277
/replace=""
/note="A substituted in clone: XX-540F24"
12277
/replace="c"

variation 12290..12291
/note="deleted in clone: XX-540F24"
/replace="a"
variation 12312
/note="G substituted in clone: XX-540F24"
/replace="a"
variation 13362
/note="G substituted in clone: XX-1Bcos"
/replace="a"
variation 13768
/note="T substituted in clone: XX-540F24"
/replace="g"
variation 14138
/note="T substituted in clone: XX-540F24"
/replace="a"
variation 14664
/note="C substituted in clone: XX-540F24"
/replace="t"
variation 15355
/note="deleted in clone: XX-1Bcos"
/replace="a"
variation 15447..15448
/note="deleted in clone: XX-1Bcos"
/replace="a"
variation 15481..15483
/note="deleted in clone: XX-540F24"
/replace="a"
variation 15888
/note="A substituted in clone: XX-540F24"
/replace="g"
variation 15961
/note="C substituted in clone: XX-540F24"
/replace="t"
variation 16252
/note="C substituted in clone: XX-540F24"
/replace="t"
variation 16334
/note="A substituted in clone: XX-1Bcos"
/replace="g"
variation 16684
/note="C substituted in clone: XX-540F24"
/replace="a"
variation 16793
/note="G substituted in clone: XX-540F24"
/replace="a"
variation 16817
/note="C substituted in clone: XX-540F24"
/replace="g"
variation 17129
/note="C substituted in clone: XX-540F24"
/replace="t"
variation 17489
/note="C substituted in clone: XX-540F24"
/replace="g"
misc_feature 18712..18717
/note="single stranded/single chemistry region"
misc_feature 21509..21686
/note="single stranded/single chemistry region"
variation 21628
/note="A substituted in clone: XX-540F24"
/replace="g"
variation 21652
/note="G substituted in clone: XX-540F24"
/replace="a"
misc_feature 22356..22558
/note="single stranded/single chemistry region"
misc_feature 22529
/note="single clone coverage"
misc_feature 22559..22586
/note="single clone coverage"
misc_feature 22606
/note="low quality region"
unure 22606

Query Match 9.2%; Score 461.6; DB 9; Length 32799;
Best Local Similarity 63.9%; Pred. No. 6,3e-110;
Matches 838; Conservative 0; Mismatches 444; Indels 30; Gaps 8;
QY 325 GGGGTGTCATGATCATTCTTTGACGGAG 384
DB 6412 GGTATTAATCTATGATATCTCTTTGAGGAG 6353
QY 385 TCTGCTCTGCGCCAGGTGAGTGCAGTGGGAGTCTGCTCACTGCAAGCTCCG 444
DB 6352 TCTGCTCTGCTGCGCCAGGTGAGTGCAGTGGGAGTCTGCTCACTGCAAGCTCCG 6293
QY 445 CTTCCGGGTTTCAAGCATCTCTGCTGAGCTCCCAAGTGTGAGTCAAGCTGAG 504
DB 6292 CTTCCGGGTTTCAAGCATCTCTGCTGAGCTCCCAAGTGTGAGTCAAGCTGAG 6233
QY 505 CCGCCTACGCGCGGTATTTTTTTTGTATTTTGTAGAGCGGGTTTCAAGCTGTT 564
DB 6232 CCGCCTACGCGCGGTATTTTTTTTGTATTTTGTAGAGCGGGTTTCAAGCTGTT 6173
QY 565 AGCGGATGCTCTGATCTCTGAGCTGATCCGCCCTCGGCTCCCAAGTGC 624
DB 6172 TTCAGGATGCTCTGATCTCTGAGTCAATGATCCAGCTCGGCTCCCAAGTGC 6113
QY 625 TGGGATTACAGGGGTGAGCGCGCGCGCGCATGATCTTCTTGACTATGCTATG 684
DB 6112 TGGGATTACAGGGGTGAGCGCGCGCGCATGATCTTCTTGACTATGCTATG 6053
QY 685 TGCAAGTACTTAAGCCATCAGACTCTTAAATATGAGTTGGGCCAGGACC 744
DB 6052 ATTAAATTAAGTGTGCGCGCTTTAGA-----AACTGGACACTTGGCCAGGACA 5999
QY 745 GTGCTCATGCTCTGATATTCAGACTTTTGGAGGAGAGTGGTGAATCACTTGGC 804
DB 5998 GTGCTCATGCTCTGATATTCAGACTTTTGGAGTGGTGAATCACTTGGC 5941
QY 805 CAGAGTTTGAACAAGCTGCGCAACATGTTGAACTGTCTTAAATTAATTAATTA 864
DB 5940 CAGAGTTTGAACAAGCTGCGCAACATGTTGAACTGTCTTAAATTAATTAATTA 5891
QY 865 AAAAAAAAAAATCAGCGGCTGTGTTGGGCAACCTGTATCCAGCTATGCTGA 924
DB 5890 --AATATCAAAAATTAAGTGGAGTGTGGCCATCTGTATCCCAAACTTGGGA 5833
QY 925 GGTGAGGCAAGAGTCACTTGAACCTGAGGCGGAGTTTGCAGTGGCCCAAGATCAG 984
DB 5832 GGTGAGGCAAGAGTCACTTGAACCTGAGGCGGAGTTTGCAGTGGCCCAAGATG 5773
QY 985 ATCAGCGCTCCAGCTGCGGCAAGAGCAAGTCTGTCAATTAATTAATTAACA 1044
DB 5772 GCTTGTCACTCCAGCTT--GTTACAGAGCAAGCTT-----TAAAGAAAGAGAG 5720
QY 1045 AAGCAACAGCAGTTTGTGTACTTATATCTTAAATTAATTAATTAATTAACA 1104
DB 5719 AAGCAACAGCAGTTTGTGTACTTATATCTTAAATTAATTAATTAATTAACA 5660
QY 1105 AGAGCAGACTGAATTAAGAAATTAATGAGCAAGCTTAAAGTATTTTGAACA 1164
DB 5659 AAGCAACAGCAGTTTGTGTACTTATATCTTAAATTAATTAATTAATTAACA 5601
QY 1165 ATCATTGAACTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAACA 1224
DB 5600 AAGCAACAGCAGTTTGTGTACTTATATCTTAAATTAATTAATTAATTAACA 5541
QY 1225 GGGAGCAGAAACACTGTGG--CCAAAGAAACAAACAGGCTCTTAAGACAGCAGG 1283
DB 5540 AGGAGAAAGAAAG 5481
QY 1284 AGGAGAAAG 1343
DB 5480 GGGAGAAAG 5421
QY 1344 GAGACAG 1402

```

Db      5420 TGGAAAGAAAGAAAGCAAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5361
QY      1403 CCGAGATATTGAGAGAGACTCAGAAAGATAGCCAGAGAGAAACCAACAGAGATGAGAAAG 1462
Db      5360 AAGAAAGAAATGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAG 5301
QY      1463 GACTCTGAGAAAAAACAGAGACAAAGATGAGAAAGAGATATCGAGGGTGAACACAGAC 1522
Db      5300 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5241
QY      1523 TGGTGAATGAGCAAAATGCAAGAAAGAAAGCAAGCAATCCAGCGCCAGAAATAGTGAC 1582
Db      5240 AGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAG 5181
QY      1583 CCAAGATTGCTGAGAGCCAGATCCTTAAGCGCTGGGGAGAGCGAGAGAGG 1634
Db      5180 GGAAGAGAGAGAAAGAAAGAAAGAGAGAGAAAGAAAGAAAGAAAGAG 5129

```

```

RESULT 11
AC149102/c 190603 bp DNA linear HTG 13-MAY-2004
LOCUS AC149102
DEFINITION Papio anubis clone RP41-113H15, WORKING DRAFT SEQUENCE, 17
unordered pieces.
ACCESSION AC149102
VERSION AC149102.1 GI:47155599
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Papio anubis (Olive baboon)
ORGANISM Papio anubis

```

```

REFERENCE
AUTHORS Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
Boutfair,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
Coleman,H., Daki,N., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B.,
Idol,J.R., Jones,C., Karlins,E., Kim,H., Kwong,P., Latic,P.,
Larsen,S., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margillies,E.H., Masiello,C., Maekert,B., McDowell,U.,
Mullikin,J.C., Paguitigan,C., Portnoy,M.E., Prasad,A., Puri,O.,
Reddy-Dugue,N., Schandler,K., Schaefer,M.G., Shah,K., Sison,C.,
Stantirpou,S., Thomas,J.W., Thomas,J., Tsipouri,V., Vogt,J.L.,
Wetherby,K.D., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative

```

```

TITLE
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190603)
AUTHORS Green,E.D.
JOURNAL Direct Submission

```

```

COMMENT Submitted (13-MAY-2004) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zooemhgrl.nih.gov
----- Project Information
Center project name: gzf
Center clone name: 113H15
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 178234 bases at least Q40
Consensus quality: 180810 bases at least Q40
Consensus quality: 182566 bases at least Q20
Insert size: 223000; agarose-ef
Insert size: 189003; sum-of-contigs
Quality coverage: 10.1ix in Q20 bases; agarose-fp
Quality coverage: 11.9ix in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently

```

FEATURES

source

```

* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2285: contig of 2285 bp in length
2385: gap of unknown length
2386: contig of 3769 bp in length
6154: gap of unknown length
6155: gap of unknown length
6254: gap of unknown length
6255: contig of 4391 bp in length
10645: gap of unknown length
10646: gap of unknown length
10746: contig of 3853 bp in length
14608: gap of unknown length
14609: gap of unknown length
14709: contig of 6769 bp in length
21477: gap of unknown length
21577: gap of unknown length
21578: contig of 8249 bp in length
29827: gap of unknown length
29926: gap of unknown length
34784: contig of 4858 bp in length
34785: gap of unknown length
34884: gap of unknown length
40439: contig of 5555 bp in length
40439: gap of unknown length
40440: gap of unknown length
40539: gap of unknown length
40540: contig of 7356 bp in length
47895: gap of 7356 bp in length
47995: gap of unknown length
47996: contig of 15311 bp in length
63306: gap of unknown length
63406: gap of unknown length
63407: contig of 17150 bp in length
80556: gap of unknown length
80557: gap of unknown length
80566: gap of unknown length
80567: contig of 8113 bp in length
88770: gap of unknown length
88779: gap of unknown length
97862: gap of 8993 bp in length
97863: gap of unknown length
97962: gap of unknown length
109400: contig of 11438 bp in length
109501: gap of unknown length
109502: gap of unknown length
133884: contig of 24384 bp in length
133885: gap of unknown length
133985: gap of unknown length
166528: contig of 32644 bp in length
166629: gap of unknown length
166729: contig of 23875 bp in length.
Location/Qualifiers
1..190603
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-113H15"
/clone_lib="RP41"
/note="BAC resource: http://bacpac.choxi.org/"
1..2285
/note="assembly_fragment"
2386..6154
/note="assembly_fragment"
6255..10645
/note="assembly_fragment"
10746..14608
/note="assembly_fragment"
14709..21477
/note="assembly_fragment"
21578..29826
/note="assembly_fragment"
29927..34784
/note="assembly_fragment"
34885..40439
/note="assembly_fragment"
40539..40540
/note="assembly_fragment"
vector_side:right"
vector_end:SP6
vector_side:right"
40540..47895
/note="assembly_fragment"
47996..63306
/note="assembly_fragment"
63407..80556
/note="assembly_fragment"
80657..88769

```


Db	91126	CTCCAGGTTTCAGGCCATTCTCTGCTCAGCCTCCGAGTACCTGGAGCTACAGGCCGCC	91185			
OY	506	CGCCACTACGCCCGGCTAAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGTTTTA	565			
Db	91186	CGCCACCAACGCCCGGCTAAATTTTTTTGTATTTTTAGTAGAGACGGGGTTTCACCGGTTA	91245			
OY	566	GCCGGGATGGCTTGATCTCTCTGACCTCTGATTCGCGCCGCTCCGCTCCCAAGTCT	625			
Db	91246	GCCAGGATGGCTTGATCTCTCTGACCTCTGATTCGCGCGCTCCGCTCCCAAGTCT	91305			
OY	626	GGGATTCAGAGCGTGAGCCACCGCGGCCCGGACATGATCATCTTCTTACATGCTGATGT	685			
Db	91306	GGGATTCAGAGCGTGAGCCACCGCGGCCCGGACATGATCATCTTCTTACATGCTGATGT	91364			
OY	686	GACAAGTACCTAAAGCCATGAGACTCTAACCTTTAAATATGACGTTTGGAGCCAGCACCG	745			
Db	91365	GCTCATCTAAATGAAGAAATTTATGACAGGGGTGTTTAAACCTAGGACGGGATGGGATCA	91424			
OY	746	TGCTCATGCTCTGTAATTCACGACTTTTGGAGGACAGAGTGGGTAAATCACTTAGGCC	805			
Db	91425	TGCTCATGCTCTGTAATTCACGACTTTTGGAGGACAGAGTGGGTAAATCACTTAGGCC	91484			
OY	806	AGGAGTTTGAAGCCAGCGCTGGCCAAACATGGTGAACCTGTCTTAACTTAAAAA	865			
Db	91485	CGGAGTTTGAGAACCAAGTCTGGCCAAAGTGTCAAAACCCCATCTTACT-----	91532			
OY	866	AAAAAAAAAAAAATCACCCGGGAGTGTCTGGGGGACACCTGTATCCAGCTATGTGAG	925			
Db	91533	AAAAAAAAAAAAATCACCCGGGAGTGTCTGGGGGACACCTGTATCCAGCTATGTGAG	91592			
OY	926	GCTGAGGACGAGAGTCACTTTGAACTCTGAGGCGGAGGTTGCAATGGGCGGATCACA	985			
Db	91593	TCTGAGGATGAGATCACTTTGAACTCTGAGGCGGAGGTTGCAATGGGCGGATCACA	91652			
OY	986	TCACCGGCTTCACCGCTGGGGGACAGAGCAAGACTGTCTCAATTAATTAATAACA	1045			
Db	91653	CCACTGTACTCCAGCTGGGGGACAGAGCAAGACTGTCTCAATTAATTAATAATAA	91712			
OY	1046	ACGAA 1050				
Db	91713	AATTA 91717				
RESULT 13						
HS95C20/c						
LOCUS	HS95C20	138849 bp	DNA linear PRI 05-MAR-2003			
DEFINITION	Human DNA sequence from clone R1-95C20 on chromosome Xp11.3-11.4, complete sequence.					
ACCESSION	Z97181					
VERSION	Z97181.1	GI:3059064				
KEYWORDS	HTG.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
AUTHORS	1 (bases 1 to 138849)					
TITLE	McDougal1.R.					
JOURNAL	Direct Submission					
COMMENT	Submitted (05-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Apr 18, 1998 this sequence version replaced gi:2222679.					
----- Genome Center						
Center: Wellcome Trust Sanger Institute						
Center code: SC						
Web site: http://www.sanger.ac.uk						
Contact: humquerry@sanger.ac.uk						

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the validation annotation may not be found in the sequence submission						

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chix> RPI-95C20 is from the library RPI-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pCYPAC2.

```

FEATURES
SOURCE
location/Qualifiers
1..138849
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPC1P704C2095"
/db_xref="taxon:9606"
/chromosome="X"
/map="p11.3-11.4"
/clone="RP1-85C20"
/clone_id="RPC1-1"

```

ORIGIN

Query Match	9.1%	Score	455	DB	9	Length	138849
Best Local Similarity	77.6%	Pred. No.	4.1e-108				
Matches	583	Conservative	0	Mismatches	150	Indels	18
						Gaps	2

OY	333	TCTATATCATCTCTTGAAGACGGAGTCCTGCTC	392
Dd	122178	TCCAAACTATTATTATCTTGAAGCGGAGTCGTGCTC	122119
OY	393	TGTGCSCCAAGCGTGAGTGCAGTGGCCGGAGATCTGGGCTCACTGCBAAGCTCCGCTCCCGG	452
Dd	122118	TGTGGCCCAAGCGGGGAGTGCAGATGGCCGAATCTCGGCTCACTGCBAAGCTCCGCTCCCGG	122059
OY	453	GTTCAAGCCATCTCCTGTCTAGGCTCCCAAAGTAGCTGGGGACTAAGGGGCCCGGCACCT	512
Dd	122058	GTTCAAGCCATCTCCTGTCTAGGCTCCCGAGTAGCTGGGAGCTAAGGGGCCCAACATC	121999
OY	513	ACGCCGGGCTAATTTTTTTTGTATTTTGTAGTAGACGGGGTTTCAACGGTTTATGACCCGGGA	572
Dd	121998	ACGCCGGGCTAATTTTTTTTGTATTTTGTAGTAGAGACGGGGTTTCAACGGTGTATGACCAGGA	121939
OY	573	TGGCCTCGATCTCTBACCTCGTGCATCCGCCCGCTCGGCTCCCAAAGTGTGGAGTTA	632
Dd	121938	TGGTCTCGATCTCTBACCTCGTGCATCCGCCCGCTCGGCTCCCAAAGTGTGGAGTTA	121879
OY	633	CAGCGGTGAGCAACGGCGCGCGGGCCATGATCATCTTCTTGAACATAAGCTGATGTGACAAGT	692
Dd	121878	CAAGCGGTGAGCAACGGCGCGCGGGCCACTATTTTAATCTTAAATAGAATTAAGTTTCTACA	121819
OY	693	ACCTAAGGCATCAAGACTCTACCTCTTAAATA-----TGCAGTTTGGGGCCAGGCAACGT	746
Dd	121818	TCTAACCTTGATATTTTGGCCGTTCAAGTTCTAAACATAACAAGTTCTCCACGCGGGCGTGT	121759
OY	747	GCGTCATGCGCTGTAAATTCAGACATTTTGGAGGCAAGAGTGGTGAATCACTTGAGGCCA	806
Dd	121758	GCGTCATGCGCTGTAAATCCAGACATTTTGGAGGCGGAGATGTGCAGATCAACTGAGGTCA	121699
OY	807	GGAAGTTTGAGCACGCGCTGGCCAAATGCTGAACCTGTCTTTTACTTAAAAAAA	866

Db	121698	GGAGTTGGAGACACAGCCTGGCCAAACATGGTGAAAACCTGCTCTACT-----A	121651
Qy	867	AAAAAAAAAATCAAGCCGGGTGTGCTGGGGCACACCTGTAATCCAGCTATGCTGGAGG	926
Db	121650	AAAAATTCAAAATTTAGCTGGGCTGGTGGCGGGGGAACCTGTAGTCCAGCTACTCAAGAGG	121593
Qy	927	CTGAGGCAAGAGAGTCACTTGAACCCCTGGAGGCGAGGTTGCAGTGGGCGGAGATCAAT	986
Db	121590	CTGAGGCAAGAGAAATGCTGTGAACCTGGGAGGCAAGGTTGCCAATGAGCCGAGATTGGCG	1215311
Qy	987	CACGCGCCCTCCAGCCTGGGCGACAGAGCAAGACTGTGTCTCAATATAATATAAACA	1046
Db	121530	CACGTCACTCCAGCCTGGGTGACAGATAGAGACTCATCTCAAAAAATAATATAAAGT	121471
Qy	1047	CGAACAGAGATTGTGTACCTTAGTTATA	1077
Db	121470	AAATTAACTAGACAGCTCTGCTTGACTTAACA	121440

2222	2308	7571	7564	13999	13754
360	<800	4492	4508	9092	9112
2344	2308	1092	1054	2356	2380
1198	1180	3332	3396	2287	2380
3828	3831	1123	1054	2669	2713
3328	3300	1803	1824	924	930
3248	3300	3041	3122	4641	4650
1289	1252	16514	16912	1247	1206
399	<800	192	<800	1708	1662
4134	4173	4763	4687	6082	6133
900	895	587	<800	1904	1897
3852	3963	5086	5041	4433	4394
4463	4378	2278	2421	6851	6941
3959	3831	1336	1314	1799	1783
1990	2005	3223	3122	23865	24086
4272	4173	7004	7057		
2327	2308	971	1054		
6858	6866	2341	2421		
3056	3063	1649	1646		
6815	6866	2822	2867		
809	810	6084	6048		
4433	4378	163	<800		
6135	6032	9648	9389		
		311	<800		
		3279	3396		
		3900	3950		
		2035	2029		
		3612	3630		
		3053	3229		
		3414	3396		
		865	883		
		637	<800		
		259	<800		
		97	<800		
		1242	1314		
		270	<800		
		23	<800		

[illegible]

7582	7368	3318	3369	10618	10357
61	<800	44	<800	7321	7358
1886	1877	2371	2346	2226	2251
1453	1416	6367	6431	1370	1327
4192	4145	3577	3633	5827	5912
17721	17419	303	<800	371	<8000
7973	7948	2574	2596	9765	9439
395	<800	629	<800	5615	5530
2993	3003	2820	2850	7814	7838
646	<800	9059	9045	1345	1327
3902	3887	1553	1517	978	998
74	<800	1792	1741	2723	2775
1930	2088	7744	7870	1059	1071
73	<800	18454	18910	2640	2595
5859	5809	4671	4671	4019	4073
1198	1186	8958	9045	3118	3126
4305	4376			5282	5530
6273	6405			83	<8000
1719	1681			3648	3652
482	<800			5541	5530
504	<800			432	<8000

Db 91802 GCTCATCTAATGAAGAAATTATGACAGGGGTGTTAAATCTAGGACGAGGGATGGGCATGA 91861

QY 746 TGGCTCATGCTGTAAATTCAGACACTTTGGGAGGCGAGGTGGGTGAATCATCTTGAGGCC 805

Db 91862 TGGCTCACGGCTGTAAATCCAGACACTTTGGGAGGCGAAGGCCAATGGATCATCTGAAGTC 91921

QY 806 AGGAGTTTGAGACCGAGCTGGCCAAATGGTGAATCTGTCTTAACTAAAAA 865

Db 91922 CGGAGTTTCGAGACCAAGTCTGGCCAAAGTGCANAAACCCATCTCTACT----- 91969

QY 866 AAAAAAAAAAATCAGCCGGGTGTGTGGGGCAACCTGTAAATCCCACTATGCTGGAG 925

Db 91970 AAAAAATCAAAAAATTAAGCCAGGGCTTGTGTGTGTGTGCTGTGTATCCCACTATCTGGGAG 92029

QY 926 GCTGAGGACGAGAGTCACTTGAACCTGGAAGCGAGGTTGCAGTGGGCCGAGATCACA 985

Db 92030 TCTTAGGCAATGAGAAATCATCTTGAACCTGGGAGGTGAGGTTGCAGTGAAGCCAAAGTTACG 92089

QY 986 TCACCGGCTTCAGAGCTGGGCGACAGAGCAAGACTGTGTCTCAATTAATTAATTAACA 1045

Db 92090 CCACTGTACTCCAGCGCTGGGTGACAGAGCAAGACTCTGTCTCAAAAAAATTAATTAATTA 92149

QY 1046 ACGAA 1050

Db 92150 AATAA 92154

Search completed: February 26, 2005, 06:23:44
Job time : 21330.3 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 / Search time 2531.5 Seconds

(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271B-13_COPY_1_5000
Perfect score: 5000

Sequence: 1 gggccagagtgaggaag.....tgcatccacagtgctgctc 5000

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1980s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2002as:*

6: geneseqn2002as:*

7: geneseqn2002as:*

8: geneseqn2003as:*

9: geneseqn2003as:*

10: geneseqn2003as:*

11: geneseqn2003as:*

12: geneseqn2004as:*

13: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5000	100.0	11570	3	AAA95905 Human KLK
2	5000	100.0	11570	12	ADK52482 Human KAL
3	5000	100.0	11570	13	ADR72623 Human KAL
4	5000	100.0	11570	13	ADR72875 Human KAL
5	472	9.4	525	12	ACH80345 Human KAL
6	440.2	8.8	41454	11	ACN44942 Human KAL
7	424.4	8.5	84607	2	AAK90847 Human PAC
8	419	8.4	91352	12	ADN94799 Human PAC
9	418.4	8.4	115935	13	ADR52880 Drug Ther
10	417.6	8.4	26371	11	ACN44814 Human gen
11	417.6	8.4	91760	11	ACN44410 Human gen
12	417.4	8.3	216215	10	ADFE9167 Human MP5
13	416.6	8.3	139904	6	ABK83562 Human CDN
14	416.2	8.3	79652	10	ADG70184_3 Continuation (4 of
15	415.2	8.3	17397	4	AAK36445 Human car
16	415.2	8.3	17397	10	ADK47139 Human car
17	415.2	8.3	17397	13	ADJ08557 Human car
18	415.2	8.3	19334	4	AAK36443 Human car
19	415.2	8.3	19334	10	ADK47137 Human car
20	415.2	8.3	19334	13	ADJ08555 Human car

21	415.2	8.3	19345	4	AAK36444 Human car
22	415.2	8.3	19345	10	ADK47138 Human car
23	415.2	8.3	19345	13	ADJ08556 Human car
24	415.2	8.3	29001	12	ADK85575 Human ztn
25	414.8	8.3	12970	4	AAK05001 Human tes
26	414.8	8.3	12970	4	ABL78894 Human tes
27	414.6	8.3	74037	6	ABK94412 DNA encod
28	414.6	8.3	11084	12	ADJ018608 Human sof
29	414.2	8.3	212221	11	ACN44598 Human gen
30	410.4	8.2	10445	4	AAK03652 Human rep
31	410.6	8.2	50000	9	ADK16926 Human DYX
32	408.2	8.2	240000	8	ACD13446 Human DNA
33	407.6	8.2	196686	11	ACN44170 Human gen
34	407	8.1	32460	12	ADK07738 Nucleotid
35	406.2	8.1	12970	4	AAK05001 Human rep
36	406.2	8.1	12970	4	ABL78894 Human tes
37	406	8.1	32224	4	AAK89966 Human dig
38	405.8	8.1	4433	4	AAK04142 Human rep
39	405.8	8.1	4433	5	ABK18400 Human nar
40	405.8	8.1	4433	5	AAK40516 DNA encod
41	405.8	8.1	4433	11	ADK09722 Human pro
42	405.8	8.1	126001	12	ADK77123 Human PAZ
43	405	8.1	32767	8	AAK56075 Human SNL
44	405	8.1	32767	9	ADK02437 Human SNL
45	405	8.1	32767	10	ADK72176 Human SNL

ALIGNMENTS

RESULT 1
ID AAA95905 standard; DNA; 11570 BP.

XX AAA95905;

DT 02-FEB-2001 (first entry)

XX Human KLK-L2 gene.

XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;
KW kallikrein-like protein; serine protease; cytosolic; cancer;
KM prostate cancer; ds.

XX Homo sapiens.

OS WO200053776-A2.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-CN000258.

XX PR 11-MAR-1999; 99US-0124260P.

PR 01-APR-1999; 99US-0127386P.

PR 21-JUL-1999; 99US-0144919P.

XX (MOUN) MOUNT SINAI HOSPITAL.

PI Yousef GM, Diamandis EP;
WPI; 2000-587440/55.
DR P-FSDB; AAK81296.

DR New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L
protein mediated disorders, especially cancer.

XX Claim 1; Page 143-149; 184pp; English.

PS The present sequence is the coding sequence of the human KLK-L2 gene,
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like
CC proteins are a subgroup of the serine protease enzyme family. They
CC catalyse the selective cleavage of specific polypeptide precursors to
CC release peptides with potent biological activity. Nucleic acids encoding

CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and KLK-L6 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 5000; DB 3; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 GGGCCAGAGTGAAGCAGAGAGAGGTTGAGAGCTCCCTGTCAAAAGTGCTTGAAGTC 60
Db 1 GGGCCAGAGTGAAGCAGAGAGAGGTTGAGAGCTCCCTGTCAAAAGTGCTTGAAGTC 60

Qy 61 TCCCTGCTTAAATGCAAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 61 TCCCTGCTTAAATGCAAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120

Qy 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

Qy 181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAGAG 240

Qy 241 ACACAAATGAGAGACACAGAGGTGTAAGAAAGAGAGATTAACAGAGTCCAGATAC 300
Db 241 ACACAAATGAGAGACACAGAGGTGTAAGAAAGAGAGATTAACAGAGTCCAGATAC 300

Qy 301 AAAGGGGACGAAGCAGTTTTCAGGGTGCTGATGATCACTCTTTTCTTTTCTTTT 360
Db 301 AAAGGGGACGAAGCAGTTTTCAGGGTGCTGATGATCACTCTTTTCTTTTCTTTT 360

Qy 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 420
Db 361 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 420

Qy 421 GATCTGAGCTCACTGCAAGCTCCGCTCCGAGGTTACAGGCATTTCTTCTGCTCAG 480
Db 421 GATCTGAGCTCACTGCAAGCTCCGCTCCGAGGTTACAGGCATTTCTTCTGCTCAG 480

Qy 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCGCTAATTTTTTGAATTTT 540
Db 481 CCAAGTAGCTGGGACCTACAGGCGCCGCACTACGCGCGCTAATTTTTTGAATTTT 540

Qy 541 GTAGAGACGGGATTTCACGCTTTTACCGGATGAGCTCGATCTCTGACCTGTGATCC 600
Db 541 GTAGAGACGGGATTTCACGCTTTTACCGGATGAGCTCGATCTCTGACCTGTGATCC 600

Qy 601 GCCCGCTGGGCTCCCAAAAGTCTGAGATTACAGGCGTGAAGCCAGCGCGCCAGT 660
Db 601 GCCCGCTGGGCTCCCAAAAGTCTGAGATTACAGGCGTGAAGCCAGCGCGCCAGT 660

Qy 661 ATCATCTTCTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
Db 661 ATCATCTTCTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 720

Qy 721 AATATGAGTTTGGGACAGGACCGTGGCTCATGCTGTAAATTCAGACCTTTGGAGGC 780
Db 721 AATATGAGTTTGGGACAGGACCGTGGCTCATGCTGTAAATTCAGACCTTTGGAGGC 780

Qy 781 AGAGGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCAGCTGGCAATGATGAAA 840
Db 781 AGAGGTGGGTGAATCACTTGAGGCGCAGAGTTTGAAGCAGCTGGCAATGATGAAA 840

Qy 841 CTCTGTCTTAACTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 900
Db 841 CTCTGTCTTAACTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 900
```

```
Qy 901 ACTGTAAATCCAGCTATCTGAGGCTGAGGACAGAGATCACTTGAACCTTGAGGCG 960
Db 901 ACTGTAAATCCAGCTATCTGAGGCTGAGGACAGAGATCACTTGAACCTTGAGGCG 960

Qy 961 GAGGTTCAGTGGCCGAGATCAATCACCGCTTCACGCTGGGCGACAGACAGACT 1020
Db 961 GAGGTTCAGTGGCCGAGATCAATCACCGCTTCACGCTGGGCGACAGACAGACT 1020

Qy 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
Db 1021 CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080

Qy 1081 AAAAAAATATCTGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 AAAAAAATATCTGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140

Qy 1141 AGAATCTTAAGTATTTTGAATAATCATTCAGAACCTTTTAAATTAATTAATTAAT 1200
Db 1141 AGAATCTTAAGTATTTTGAATAATCATTCAGAACCTTTTAAATTAATTAATTAAT 1200

Qy 1201 GGCATTGAAGAAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 GGCATTGAAGAAGCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260

Qy 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320

Qy 1321 AAAAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AAAAAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380

Qy 1381 GAGAGGGGGTGAAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 GAGAGGGGGTGAAGAGAGAGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAG 1440

Qy 1441 GAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500

Qy 1501 AGTATCGAGGTTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAGAGAGAG 1560
Db 1501 AGTATCGAGGTTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAGAGAGAG 1560

Qy 1561 TCCAGGCGCAAGAAATGAGACCAAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TCCAGGCGCAAGAAATGAGACCAAGTTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

Qy 1621 AGGCAAGGAGAGGAGCTGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 AGGCAAGGAGAGGAGCTGGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680

Qy 1681 GTAGGAGTGAATTCGAGCTGGGTTGGGGGTGCTCTGAGGGGTGAGATAGGGGAGCA 1740
Db 1681 GTAGGAGTGAATTCGAGCTGGGTTGGGGGTGCTCTGAGGGGTGAGATAGGGGAGCA 1740

Qy 1741 GTAGGAGTGAATTCGAGCTGGGTTGGGGGTGCTCTGAGGGGTGAGATAGGGGAGCA 1800
Db 1741 GTAGGAGTGAATTCGAGCTGGGTTGGGGGTGCTCTGAGGGGTGAGATAGGGGAGCA 1800

Qy 1801 CAGTGGGTGTTAATCTAGGCTCCGATGAGGACCTCATTTGCCGAGAAATGATGCCAG 1860
Db 1801 CAGTGGGTGTTAATCTAGGCTCCGATGAGGACCTCATTTGCCGAGAAATGATGCCAG 1860

Qy 1861 AGGCAAGGCTGAGAGAGTCTGCGCTGAGCTGGAGCAAAATCCCACTTCTGAG 1920
Db 1861 AGGCAAGGCTGAGAGAGTCTGCGCTGAGCTGGAGCAAAATCCCACTTCTGAG 1920

Qy 1921 GGGCAAGGCTGAGAGAGTCTGCGCTGAGCTGGAGCAAAATCCCACTTCTGAG 1980
Db 1921 GGGCAAGGCTGAGAGAGTCTGCGCTGAGCTGGAGCAAAATCCCACTTCTGAG 1980

Qy 1981 GTGCGTCTGACCACTTCTCTGTCTCTGCTCTTCTGCTCTGAGAGCTGTAAGA 2040
```

```

1981 GTGGGTCTGCAACCAACATCTTCTGTGTCCCTCTTGCCCTGTGTGAGGCTGCTAGA 2040
Qy CTCTATCTTCTGAAATCTATAGTCCCTGAGCTTCAAGCAGATGCCATGCTGACCTGTC 2100
Db CTCTATCTTCTGAAATCTATAGTCCCTGAGCTTCAAGCAGATGCCATGCTGACCTGTC 2100
Qy CTGTGGTCTCTCTCTACCTGAGGAAATTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db CTGTGGTCTCTCTCTACCTGAGGAAATTAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Qy GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGATGCGAGCGCC 2220
Db GCTCCCGGATCGCTGGGCTCCCAACCTCTGACATTCCTCCATCCAGATGCGAGCGCC 2220
Qy ATGCTACAGCAAAACCCCTGATGAGTGGTCTGTGCTCTGATCAACAGCTTGCTT 2280
Db ATGCTACAGCAAAACCCCTGATGAGTGGTCTGTGCTCTGATCAACAGCTTGCTT 2280
Qy CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Db CTGGGGGTCAAGGTAAACAGAACTCTGGGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Qy CTGTGCGGCACTAGAGGCGCTGTCCCTGGGAACTGTGTGAGCTGGGCACTGACCTCGG 2400
Db CTGTGCGGCACTAGAGGCGCTGTCCCTGGGAACTGTGTGAGCTGGGCACTGACCTCGG 2400
Qy GAGTGGGTGAATGATGATCTGTCTGTATCTTGTGTGTGATGATGATGATGATGATGATG 2460
Db GAGTGGGTGAATGATGATCTGTCTGTATCTTGTGTGTGATGATGATGATGATGATGATG 2460
Qy GACTGCCACGGGTGTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db GACTGCCACGGGTGTGTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Qy CAGGTGGCACTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db CAGGTGGCACTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Qy ATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
Db ATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
Qy GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Db GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2700
Qy TCCCTGAGGCTCCCGGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Db TCCCTGAGGCTCCCGGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
Qy GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2820
Db GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2820
Qy TTTGTGACCGGTGATCACTGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Db TTTGTGACCGGTGATCACTGAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Qy TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Db TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
Qy TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Db TTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3000
Qy CAGGTGATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
Db CAGGTGATGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
Qy TGTGACACACGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Db TGTGACACACGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
3061 TGTGACACACGCGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120
Qy GCTAGGCTGCGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
Db GCTAGGCTGCGCGGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3180
Qy GAGGCAACATGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Db GAGGCAACATGAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3240
Qy GCGTGTCTTGTGGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Db GCGTGTCTTGTGGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3300
Qy GAAACAGGCGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Db GAAACAGGCGGCGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3360
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3420
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3420
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3480
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3480
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3540
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3540
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3600
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3600
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3660
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3660
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3720
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3720
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3780
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3780
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3840
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3840
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3900
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3900
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3960
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 3960
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4020
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4020
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4080
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4080
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4140
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4140
Qy GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4200
Db GCGGATCACTGAGGTGCGGAGATCGAGGCTGACCAATGAGAGAACCCGCT 4200
4141 GTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200
Db GTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4200
4141 GTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
Db GTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
```

QY	4201	TCAGTAGTATCCGGCCACCTCGGCGCTCCCAAGATGCTGGGGTTTACAGAGGTGAACAACCGC	4280
Db	4201	TCAGGTGATTCGGCCACCTCGGCGCTCCCAAGATGCTGGGGTTTACAGAGGTGAACAACCGC	4280
QY	4261	CCCCAGCCCAAAGTCAGAGCTCTTTATATGAGAGCTCTTAACATGTAACTGTAACTCTGACCTGAGCC	4320
Db	4261	CCCCAGCCCAAAGTCAGAGCTCTTTATATGAGAGCTCTTAACATGTAACTGTAACTCTGACCTGAGCC	4320
QY	4321	CTAACTAATGTCATTTCCAAACCCCTTCTGCTCCAGCCTTGACCCCACTCACTGAGGCC	4380
Db	4321	CTAACTAATGTCATTTCCAAACCCCTTCTGCTCCAGCCTTGACCCCACTCACTGAGGCC	4380
QY	4381	TGACCCCACTCTTGAGAGCCAGTTCACATCCCTTAAAGCCCTGCTCCCTCCCACTCCAG	4440
Db	4381	TGACCCCACTCTTGAGAGCCAGTTCACATCCCTTAAAGCCCTGCTCCCTCCCACTCCAG	4440
QY	4441	GCTCCAGCCCCCAGAGCTTTTGAGCACTAACCCCTGAGCTGTGCACAGAAATCCTGTACCCAAAT	4500
Db	4441	GCTCCAGCCCCCAGAGCTTTTGAGCACTAACCCCTGAGCTGTGCACAGAAATCCTGTACCCAAAT	4500
QY	4501	TTTAACTCTCAGATGTAGTTCTAAGCCAAATTCAGAAATCTGTAGAGTCCAATTAAAGTCCA	4560
Db	4501	TTTAACTCTCAGATGTAGTTCTAAGCCAAATTCAGAAATCTGTAGAGTCCAATTAAAGTCCA	4560
QY	4561	GTAACCCCTACAGAGCCCTGGGCTCGTCTTGAGGCTTGAGGCTGGAGGCTGAGAGGTGCA	4620
Db	4561	GTAACCCCTACAGAGCCCTGGGCTCGTCTTGAGGCTTGAGGCTGGAGGCTGAGAGGTGCA	4620
QY	4621	CTCTTATTTCTCCAGAGCCCTGCGCCCTGCGCCCTCAGCATGTGAGACACCACTCTTAGCT	4680
Db	4621	CTCTTATTTCTCCAGAGCCCTGCGCCCTGCGCCCTCAGCATGTGAGACACCACTCTTAGCT	4680
QY	4681	GCTCTGCGCTCTTGAGTCTGAATCCCAACCCCAAGCCCAAGCCCGGCTCTGAGGCCCGGCC	4740
Db	4681	GCTCTGCGCTCTTGAGTCTGAATCCCAACCCCAAGCCCAAGCCCGGCTCTGAGGCCCGGCC	4740
QY	4741	CAACCCATTTTCCGTTCCCAAGAGATGTTCTCGCAACATATGATTTTCTGTGACCAAC	4800
Db	4741	CAACCCATTTTCCGTTCCCAAGAGATGTTCTCGCAACATATGATTTTCTGTGACCAAC	4800
QY	4801	CCTCTTAAACACCGTGCCCTCTGGAGACAACAAGACTTGGAGACTGGAGCCGGGGAGAGACG	4860
Db	4801	CCTCTTAAACACCGTGCCCTCTGGAGACAACAAGACTTGGAGACTGGAGCCGGGGAGAGACG	4860
QY	4861	CCCCGTCGAGATGACAGACAGACGCGGCATCATATGATTCGAGCTGGGATTTGCAACACC	4920
Db	4861	CCCCGTCGAGATGACAGACAGACGCGGCATCATATGATTCGAGCTGGGATTTGCAACACC	4920
QY	4921	AGCCGTGGCAGAGCCGCGCTGTGTGCTAAGGCCCAACAGCTCTAACCTGCGGGGCGGGTGTGG	4980
Db	4921	AGCCGTGGCAGAGCCGCGCTGTGTGCTAAGGCCCAACAGCTCTAACCTGCGGGGCGGGTGTGG	4980
QY	4981	TGCATCCACAGTGGCTGCTC 5000	
Db	4981	TGCATCCACAGTGGCTGCTC 5000	
RESULT 2			
ADK52482			
ID	ADK52482 standard; DNA; 11570 BP.		
XX	ADK52482;		
XX			
DT	03-JUN-2004 (first entry)		
XX			
DE	Human kallikrein 5 encoding sequence.		
XX			
KW	kallikrein 5; cancer; Cytostatic; Immunostimulant; ds.		
OS	Homo sapiens.		
XX			
PN	WO2004021008-A2.		

Query Match	100.0%	Score 5000:	DB 12;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 5000;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0
1	GGGCCAGAGTGAAGGCAAGAGAGAGTTCCTCTGCAAGTGGCTTGAGTC	60		
1	GGGCCAGAGTGAAGGCAAGAGAGAGTTCCTCTGCAAGTGGCTTGAGTC	60		
61	TCCCTTGGCTTAAATGCAGGGAGAGGGGCAAGAAACACAGGGAAGAGAGGGTGGGG	120		
61	TCCCTTGGCTTAAATGCAGGGAGAGGGGCAAGAAACACAGGGAAGAGAGGGTGGGG	120		
121	AAGAAAGAGAGAGAGAGAGAGAGAGAGATTAACAACATACAGAAACACAGAGAGAACAC	180		
121	AAGAAAGAGAGAGAGAGAGAGAGAGAGAGATTAACAACATACAGAAACACAGAGAGAACAC	180		
181	ACAGAGAGCTGGGACACAGAGGACACACAGAGTCAGAGAGAAAGAAAGATAGAGAAAG	240		
181	ACAGAGAGCTGGGACACAGAGGACACACAGAGTCAGAGAGAAAGAAAGATAGAGAAAG	240		
241	ACACAAATGAGACACACAGAGGTGTAAGAAAGAGATTAACAGAGTCCCAATACACGC	300		
241	ACACAAATGAGACACACAGAGGTGTAAGAAAGAGATTAACAGAGTCCCAATACACGC	300		
301	AAAGGGGACAGAGACACAGTTTTCAGGGTGGTGTATGATCATCTCTTTT	360		
301	AAAGGGGACAGAGACACAGTTTTCAGGGTGGTGTATGATCATCTCTTTT	360		
361	TTTTTTTTTTTTTTGAGACGAGTCTTGCTCTGTGCGCCAGGCTGAGAGTCAGTGGCG	420		
361	TTTTTTTTTTTTTTGAGACGAGTCTTGCTCTGTGCGCCAGGCTGAGAGTCAGTGGCG	420		
421	GATCTCGGCTACATGCAAGCTCCGGCTCCGGGTTACAGGCATTCCTCGCTCAGGCTC	480		
421	GATCTCGGCTACATGCAAGCTCCGGCTCCGGGTTACAGGCATTCCTCGCTCAGGCTC	480		
481	CCAAAGTACTGGAGCTACAGGCGCCCGACATACGCGCCGCTAATTTTTTTTATTTTTTA	540		
481	CCAAAGTACTGGAGCTACAGGCGCCCGACATACGCGCCGCTAATTTTTTTTATTTTTTA	540		

Db 481 CCAAGTACGTGGACACAGGCGCCGCACTACGCCCGGCTAATTTTTTGTATTTTA 540
Qy 541 GTAGAGACGGGGTTTCAACCGTTTAAGCCGGGATGGCTTCATCTCTGACCTCGTATCC 600
Db 541 GTAGAGACGGGGTTTCAACCGTTTAAGCCGGGATGGCTTCATCTCTGACCTCGTATCC 600
Qy 601 GCCCGCTCGGCTCCCAAGTGTGGGATTACAGGGGTGAGCCACCGCGCCGGCCATG 660
Db 601 GCCCGCTCGGCTCCCAAGTGTGGGATTACAGGGGTGAGCCACCGCGCCGGCCATG 660
Qy 661 ATCATCTTCTGACTATGCTGATGTGACAGTACCTTAAGCCATCAGACTTACCTTTA 720
Db 661 ATCATCTTCTGACTATGCTGATGTGACAGTACCTTAAGCCATCAGACTTACCTTTA 720
Qy 721 AATTGCACTTTGGGCGCAGGCACTGCTGATGCTGATTTCCAGACTTTTGGAGGC 780
Db 721 AATTGCACTTTGGGCGCAGGCACTGCTGATGCTGATTTCCAGACTTTTGGAGGC 780
Qy 781 AGAGGTGGGTGAATCACTTGAGGCGAGGATTTGAGACAGGCTGGCCAACTGTGTAA 840
Db 781 AGAGGTGGGTGAATCACTTGAGGCGAGGATTTGAGACAGGCTGGCCAACTGTGTAA 840
Qy 841 CTCTGTCTTAACTAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCTGTGGGCAC 900
Db 841 CTCTGTCTTAACTAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGTCTGTGGGCAC 900
Qy 901 ACCGTGAATCCAGCTATGCTGAGGCTGAGGCGACAGAGTCACTTGAACCTTGGAGCG 960
Db 901 ACCGTGAATCCAGCTATGCTGAGGCTGAGGCGACAGAGTCACTTGAACCTTGGAGCG 960
Qy 961 GAGGTTCAGTGGGCGGAGATCACATGACCGGCTCCAGGCTGGGCGCAGAGCAAGCT 1020
Db 961 GAGGTTCAGTGGGCGGAGATCACATGACCGGCTCCAGGCTGGGCGCAGAGCAAGCT 1020
Qy 1021 CTGTCTCAAAATAAATAAATAAACAACAGAGTGTGTGTACTTATATATCT 1080
Db 1021 CTGTCTCAAAATAAATAAATAAACAACAGAGTGTGTGTACTTATATATCT 1080
Qy 1081 AAAAAAAAAATGCTGTCAACAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Db 1081 AAAAAAAAAATGCTGTCAACAATAGACAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Qy 1141 AGAATCTAAAGGTATATTTGACAAATCATTCAGAACCTTTAAAAAAGAAATCACAGA 1200
Db 1141 AGAATCTAAAGGTATATTTGACAAATCATTCAGAACCTTTAAAAAAGAAATCACAGA 1200
Qy 1201 GGCATAGAAAACAGGAGGAGACAGGAGACAGAAAACCTGTGGCCCAAGAGAACAA 1260
Db 1201 GGCATAGAAAACAGGAGGAGACAGGAGACAGAAAACCTGTGGCCCAAGAGAACAA 1260
Qy 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGTGAAGTGAAGACAGACAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGTGAAGTGAAGACAGACAGAG 1320
Qy 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGTGAAGAA 1380
Db 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGTGAAGAA 1380
Qy 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAGAAAATAGCCGAGGGA 1440
Db 1381 GAGAGAGGGGTGAGAGAGACAGAGATTTGAGAGAGACTCAGAAAATAGCCGAGGGA 1440
Qy 1441 GAACCAACAGAGATGGAAGAACTCTGAGAAAAAACACAGAGACAAAGATGAAAGAG 1500
Db 1441 GAACCAACAGAGATGGAAGAACTCTGAGAAAAAACACAGAGACAAAGATGAAAGAG 1500
Qy 1501 AGTATCGAGGTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
Db 1501 AGTATCGAGGTGAACAGACAGTGTGGAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560
Qy 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTTGTTGAGAGCCAGATCTCTTAAGCTGGGG 1620
Db 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTTGTTGAGAGCCAGATCTCTTAAGCTGGGG 1620

Qy 1621 AGGACGGAGAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCTCCGGGACAGGGAG 1680
Db 1621 AGGACGGAGAGGGGCTGGGCTTCCGGAGACCCCTCCCAATTCTCCGGGACAGGGAG 1680
Qy 1681 GTAGGAGTGAACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
Db 1681 GTAGGAGTGAACATTCGGACTGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
Qy 1741 GAGAGACTAATTCCTAAGGCCGATAGGCACTTCATTTGCCCGGAAATGTGCCAGGAG 1800
Db 1741 GAGAGACTAATTCCTAAGGCCGATAGGCACTTCATTTGCCCGGAAATGTGCCAGGAG 1800
Qy 1801 CAGTGGGTGTTAATCTCAGGCGCCGGTGCAGAGCCGAGAGAGAGAGAGAGAGGAGCA 1860
Db 1801 CAGTGGGTGTTAATCTCAGGCGCCGGTGCAGAGCCGAGAGAGAGAGAGAGAGGAGCA 1860
Qy 1861 AGGACAGAGCCTGAGAGATGCGGCTGAGCTGGAGCAAAATCCCACTTACTCTGG 1920
Db 1861 AGGACAGAGCCTGAGAGATGCGGCTGAGCTGGAGCAAAATCCCACTTACTCTGG 1920
Qy 1921 GGGACAGGGCAAGTGAACCTGAGAGGTGCTCAGCAGGCAAGGAGAGAGAGTGTCT 1980
Db 1921 GGGACAGGGCAAGTGAACCTGAGAGGTGCTCAGCAGGCAAGGAGAGAGAGTGTCT 1980
Qy 1981 GTGCTCTGCAACCAATCTTCTGCTCCCTCCCTGCGCCGTGCTGAGAGGCTGTAGA 2040
Db 1981 GTGCTCTGCAACCAATCTTCTGCTCCCTCCCTGCGCCGTGCTGAGAGGCTGTAGA 2040
Qy 2041 CTCCTATCTTCTGAATTTCTATAGTGCCTGGGTCTCAGCGCAGTGCAGATGGGCCCTG 2100
Db 2041 CTCCTATCTTCTGAATTTCTATAGTGCCTGGGTCTCAGCGCAGTGCAGATGGGCCCTG 2100
Qy 2101 CTGTGGTCTCTCTCTAAGGAGAAATTAAGTGAAGGAGAGAGAGAGAGAGTGGTTAAG 2160
Db 2101 CTGTGGTCTCTCTCTAAGGAGAAATTAAGTGAAGGAGAGAGAGAGAGAGTGGTTAAG 2160
Qy 2161 GCTCCCGGATGCGCTGGGCTGCCAAACCTCTGACATTCGCATCCAGGTGACGGGCC 2220
Db 2161 GCTCCCGGATGCGCTGGGCTGCCAAACCTCTGACATTCGCATCCAGGTGACGGGCC 2220
Qy 2221 ATGGCTPACAGAGAACCCCTCGATGTGGGTCTGTGCTGTGATCACAGGCTTGCT 2280
Db 2221 ATGGCTPACAGAGAACCCCTCGATGTGGGTCTGTGCTGTGATCACAGGCTTGCT 2280
Qy 2281 CTGGGGGTCAACAGTAAACAGAACTCTGGGTGGAGAGGTTGGGATTTGGAGAGACTGT 2340
Db 2281 CTGGGGGTCAACAGTAAACAGAACTCTGGGTGGAGAGGTTGGGATTTGGAGAGACTGT 2340
Qy 2341 CTCTGCGGCACTAGAGCGGCTGCTCCCTGGGAGAACTGTGAGACCTGGGCAATGACTCCGG 2400
Db 2341 CTCTGCGGCACTAGAGCGGCTGCTCCCTGGGAGAACTGTGAGACCTGGGCAATGACTCCGG 2400
Qy 2401 GACCGGGTGAATGAGTCTCTGTCTGATCTTGTGTTGGTCATGCTATGTGGCCCTGT 2460
Db 2401 GACCGGGTGAATGAGTCTCTGTCTGATCTTGTGTTGGTCATGCTATGTGGCCCTGT 2460
Qy 2461 GACTGCCAGGTGTGTGTGGGAGAGGAGATGCTTTTCCATATCAAGTGACTGTGCGG 2520
Db 2461 GACTGCCAGGTGTGTGTGGGAGAGGAGATGCTTTTCCATATCAAGTGACTGTGCGG 2520
Qy 2521 CAGGTGGCACTGACCTTTGAGGCTGTGTGTGTTGGTGTGATTTGATGATTTAAG 2580
Db 2521 CAGGTGGCACTGACCTTTGAGGCTGTGTGTGTTGGTGTGATTTGATGATTTAAG 2580
Qy 2581 ATTGTGTGTGCTTCAAGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 ATTGTGTGTGCTTCAAGCTGTGTGGTGAATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 GTGTTGGCTGTGTGTGTGTGATCTTGGCATTTGATGATGATGATGATGATGATGATGAT 2700
Db 2641 GTGTTGGCTGTGTGTGTGTGATCTTGGCATTTGATGATGATGATGATGATGATGATGAT 2700

[illegible]

Db	3781	CTGAGCCCTTCTTCGTCGTCCTTTCGGAACTTAGCTCCGCCATGAGCTTAGAC	3840
Oy	3841	CCACCTTCCTTCTCAACACAGCCCTTAGGCGAGACTTAGTGAGCCCGCTAAGGCA	3900
Db	3841	CCACTCTCTTTCCTCAACACAGCCCTTAGGCGAGACTTAGTGAGCCCGCTAAGGCA	3900
Oy	3901	CACCCCTTTGGGCGAGGCTCCACCCCTTAATCTGTGGGTACCTTCAGAACCCCTTCAA	3960
Db	3901	CACCCCTTTGGGCGAGGCTCCACCCCTTAATCTGTGGGTACCTTCAGAACCCCTTCAA	3960
Oy	3961	AGTAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGCTGTGTCTCTCCAGGCTGAG	4020
Db	3961	AGTAGAGCTTTTTTTTTTTTTTTTTTTTGGAGACAGCTGTGTCTCTCCAGGCTGAG	4020
Oy	4021	TGCAGTGGCGGATCTCGGCTCACTGCAACCTCGCTCCAGGTTCAAGTATTCGT	4080
Db	4021	TGCAGTGGCGGATCTCGGCTCACTGCAACCTCGCTCCAGGTTCAAGTATTCGT	4080
Oy	4081	GCTTCACCTTCGAGTAGCTGGGATTAAGAGTCCGCGCACACACGCTGGCTAATTTTT	4140
Db	4081	GCTTCACCTTCGAGTAGCTGGGATTAAGAGTCCGCGCACACACGCTGGCTAATTTTT	4140
Oy	4141	GTGCTTTTAGTAGAGACAGGTTTCACTTGTTGGCCAGGCTGGTCTCAATCCCAAC	4200
Db	4141	GTGCTTTTAGTAGAGACAGGTTTCACTTGTTGGCCAGGCTGGTCTCAATCCCAAC	4200
Oy	4201	TCAGGTATCCGCGCCACCTCGGCTCCCAAGTCTGGGTTACAGGCGTAGCCACGC	4260
Db	4201	TCAGGTATCCGCGCCACCTCGGCTCCCAAGTCTGGGTTACAGGCGTAGCCACGC	4260
Oy	4261	CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAACATGTAACTTGACCTTGCC	4320
Db	4261	CCCCAGCCCAAGTCAGAGCTTTTATAGAGACTTAACATGTAACTTGACCTTGCC	4320
Oy	4321	CTAATCTAATGTCATTTCCAAACCCCTTCTGTGCTCCAGCCTGACCCCACTACCTGAAGCC	4380
Db	4321	CTAATCTAATGTCATTTCCAAACCCCTTCTGTGCTCCAGCCTGACCCCACTACCTGAAGCC	4380
Oy	4381	TGACCCCACTTTTGAGACAGAGTTCATACCTTAAGGCTGTCTCCCTCCCATCCCGAG	4440
Db	4381	TGACCCCACTTTTGAGACAGAGTTCATACCTTAAGGCTGTCTCCCTCCCATCCCGAG	4440
Oy	4441	GCTCCAGCCCCCAAGCTTTGGCACTACCCCTGAGCTGTTCACAGAACTCTGTACCCAAAT	4500
Db	4441	GCTCCAGCCCCCAAGCTTTGGCACTACCCCTGAGCTGTTCACAGAACTCTGTACCCAAAT	4500
Oy	4501	TTTAACCTCACATGTAGTTCTAAGCAATTCAGGAATCTGTAGGTCOAAGTTAAGTCCA	4560
Db	4501	TTTAACCTCACATGTAGTTCTAAGCAATTCAGGAATCTGTAGGTCOAAGTTAAGTCCA	4560
Oy	4561	GTAACCCCTACCTGAGCCTGGGCTGTGCTTGAAGCTTGAAGCTTGAAGAGTGA	4620
Db	4561	GTAACCCCTACCTGAGCCTGGGCTGTGCTTGAAGCTTGAAGCTTGAAGAGTGA	4620
Oy	4621	CTCTTAATTTCTCAAGGCTGTGCTGCTCCCTCAGATGTACAGACCCCACTCTAGCT	4680
Db	4621	CTCTTAATTTCTCAAGGCTGTGCTGCTCCCTCAGATGTACAGACCCCACTCTAGCT	4680
Oy	4681	GGTGTGCGCTTGTAGTCTGAACCCACCCCAAGCCCAACCCCGCTCTGAGCCCGGC	4740
Db	4681	GGTGTGCGCTTGTAGTCTGAACCCACCCCAAGCCCAACCCCGCTCTGAGCCCGGC	4740
Oy	4741	CAACCCATTTTCCGTTTCCCAAGAGATGTCTCGCAACATGAATGATTTCCCTGAGCACCC	4800
Db	4741	CAACCCATTTTCCGTTTCCCAAGAGATGTCTCGCAACATGAATGATTTCCCTGAGCACCC	4800
Oy	4801	CCTTTAACAACGTCCTCTGTGGAGCAACAGACTCTGGAGCTGTGGGCGGGGAAGAG	4860
Db	4801	CCTTTAACAACGTCCTCTGTGGAGCAACAGACTCTGGAGCTGTGGGCGGGGAAGAG	4860
Oy	4861	CCCGGTGAGTAGACAGCAGACCGCATCAATGAGATCCAGCTGCGATATGCAACCC	4920

```
Db      4861 CCCGTCGAGATGACAGACAGCCGATCATCATGATCCGAGTATGACACACC 4920
Qy      4921 AGCGTGGCAGAGCCCGCTGTGCTAAGGCCAACCACTTACTGCGGGCGGTGTCG 4980
Db      4921 AGCGTGGCAGAGCCCGCTGTGCTAAGGCCAACCACTTACTGCGGGCGGTGTCG 4980
Qy      4981 TGCATCCACAGTGGCTGCTC 5000
Db      4981 TGCATCCACAGTGGCTGCTC 5000

RESULT 3
ADR72623
ID      ADR72623 standard; DNA; 11570 BP.
AC      ADR72623;
XX      02-DEC-2004 (first entry)
DE      Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.
XX      kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;
KM      cancer metastasis; chemotherapy; human; serine protease;
XX      chromosome 19q13.4; KLK5; ds; gene.
XX      Homo sapiens.
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      /tag= b
FT      /product= "Human renal cell carcinoma-related kallikrein
FT      5 (KLK5) protein"
FT      2221..2293
FT      /tag= a
FT      /number= 1
FT      /tag= c
FT      2294..4761
FT      /tag= d
FT      /number= 1
FT      /tag= e
FT      4762..5023
FT      /tag= f
FT      /number= 2
FT      /tag= g
FT      5024..5762
FT      /tag= h
FT      /number= 2
FT      /tag= i
FT      5763..6019
FT      /tag= j
FT      /number= 3
FT      /tag= k
FT      6020..6104
FT      /tag= l
FT      /number= 3
FT      /tag= m
FT      6105..6238
FT      /tag= n
FT      /number= 4
FT      /tag= o
FT      6239..11091
FT      /tag= p
FT      /number= 4
FT      /tag= q
FT      11092..11247
FT      /tag= r
FT      /number= 5
XX      WO2004077060-A2.
XX      10-SEP-2004.
XX      26-FEB-2004; 2004WO-CA000280.
XX      27-FEB-2003; 2003US-0451382P.
XX      (MOUN ) MOUNT SINAI HOSPITAL.
XX      Diamandis EP, Petraki CD;
XX      WPI; 2004-662077/64.
```

```
DR      P-PSDB; ADR72621.
XX      Detecting kallikrein polypeptides/polynucleotides associated with renal
PT      cell carcinoma in a patient, for diagnosing/treating the disease,
PT      comprises detecting /identifying kallikrein polypeptides/polynucleotides
PT      in a sample.
XX      Example 1; SEQ ID NO 3; 53pp; English.
XX      The invention relates to a novel method for detecting kallikrein
CC      polypeptides, or the polynucleotides encoding them, associated with renal
CC      cell carcinoma. The method comprises obtaining a sample from a patient
CC      and detecting kallikrein polypeptides, or their encoding polynucleotides,
CC      where the kallikrein polypeptides are selected from the group consisting
CC      of kallikrein 5, kallikrein 6, kallikrein 10 and kallikrein 11. The
CC      detected amounts of the kallikrein polypeptides are compared to standard
CC      amounts. The molecules of the invention demonstrate cytostatic activity,
CC      whilst the methods and kit may be useful for detecting, characterising,
CC      preventing and treating renal cell carcinoma. Furthermore, the methods
CC      may be useful for evaluating the probability of the presence of malignant
CC      or pre-malignant cells and for detecting and quantitating tumour growth
CC      and cancer metastasis. Finally, the methods may be utilised to confirm
CC      the absence or removal of all tumour tissue following surgery, cancer
CC      chemotherapy and/or radiation therapy and to monitor cancer chemotherapy
CC      and tumour reappearance. The current sequence is that of the human
CC      ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the
CC      invention which encodes a secreted serine protease and is located at
CC      chromosome 19q13.4.
SQ      Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
XX      Query Match 100.0%; Score 5000; DB 13; Length 11570;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 5000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGAGTGCTTGAGTC 60
Db      1 GGGCCAGAGTGAAGGCAAGAGAGTGAAGCTCCCTGCAAGAGTGCTTGAGTC 60
Qy      61 TCCCTGCTTAATGACGAGAGAGGAGGACAGAAAGACAGGAGAGAGAGGCTGGG 120
Db      61 TCCCTGCTTAATGACGAGAGAGGAGGAGGACAGAAAGACAGGAGAGAGAGGCTGGG 120
Qy      121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db      121 AAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Qy      181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAG 240
Db      181 ACAGAGAGCTGGGACACAGGACACACAGAGTCAAGAGAGAGAGAGAGAGAGAG 240
Qy      241 ACACAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGAGTCCAGATACAGC 300
Db      241 ACACAATGAGACACAGAGGTGTAAAGAAAGAGATTAAACAGAGTCCAGATACAGC 300
Qy      241 ACACAATGAGACACAGAGGTGTGTATGATGATCATCTCTTTT 360
Db      241 ACACAATGAGACACAGAGGTGTGTATGATGATCATCTCTTTT 360
Qy      301 AAAGGGCAGAGACAGCTTTTCAAGGTGTGTGTATGATCATCTCTTTT 360
Db      301 AAAGGGCAGAGACAGCTTTTCAAGGTGTGTGTATGATCATCTCTTTT 360
Qy      361 TTTT 420
Db      361 TTTT 420
Qy      421 GATCTGGCTCACTGACAGCTCCGCTCCGAGTTACAGGCAATTCCTCAGCTC 480
Db      421 GATCTGGCTCACTGACAGCTCCGCTCCGAGTTACAGGCAATTCCTCAGCTC 480
Qy      481 CCAAGTAGCTGGAGCTACAGGCGCCGCACTACGCGCGCTAATTTTGAATTTTA 540
Db      481 CCAAGTAGCTGGAGCTACAGGCGCCGCACTACGCGCGCTAATTTTGAATTTTA 540
Qy      541 GTAGAGAGCGGGTTTACCGTTTGAAGCGGAGTGCCTGATCTCCGAGCTGTATCC 600
Db      541 GTAGAGAGCGGGTTTACCGTTTGAAGCGGAGTGCCTGATCTCCGAGCTGTATCC 600
```


Db	541	GTAGAGAGGGGGTTTCAACCGTTTAAAGCCGGGATGGCTTCGATCTTCCTGACTCGTATCC	600
Qy	601	GCCCCGCTCGGCTCTCCAAAGTGTCTGGGATTAAGGCGGTGAGCCACCGCGCCGAGCATG	660
Db	601	GCCGCGCTCGGCTCTCCAAAGTGTCTGGGATTAAGGCGGTGAGCCACCGCGCCGAGCATG	660
Qy	661	ATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCATCAGACTTACCTTTTA	720
Db	661	ATCATCTTCTTGACTATGCTGATGTGACAAAGTACTTAAAGCATCAGACTTACCTTTTA	720
Qy	721	AATATGAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAFTTCAGACTTTGGGAGGC	780
Db	721	AATATGAGTTTGGGCGAGGACCGTGGCTCATGCTGTAAFTTCAGACTTTGGGAGGC	780
Qy	781	AGAGGTGGGTGAATCATCTTGAAGGCCAGAGATTGTAAGCACGCTGGCCAAATGATGTGAA	840
Db	781	AGAGGTGGGTGAATCATCTTGAAGGCCAGAGATTGTAAGCACGCTGGCCAAATGATGTGAA	840
Qy	841	CTCTGTCTTTTACTAAAAAATTTTTTTTTTTTTTTTTTCAAGCGGGTGTCTGTGGGCAC	900
Db	841	CTCTGTCTTTTACTAAAAAATTTTTTTTTTTTTTTTTTCAAGCGGGTGTCTGTGGGCAC	900
Qy	901	ACCTGTATCCAGTATGTCTGGAGGCTGAGAGCCAGAAAGTCACTTGAACCTGTGAGAGCG	960
Db	901	ACCTGTATCCAGTATGTCTGGAGGCTGAGAGCCAGAAAGTCACTTGAACCTGTGAGAGCG	960
Qy	961	GAGGTGAGTGGGCGGAGATCACATCAACCGCTCTCAAGCTGTGGGCGACAGAGCAAGCT	1020
Db	961	GAGGTGAGTGGGCGGAGATCACATCAACCGCTCTCAAGCTGTGGGCGACAGAGCAAGCT	1020
Qy	1021	CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT	1080
Db	1021	CTGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATCT	1080
Qy	1081	AAAAAATATGCTGTCAACAAATGACACAAAGTAAATTAAGAAATTAATTAATTAATGCGCCA	1140
Db	1081	AAAAAATATGCTGTCAACAAATGACACAAAGTAAATTAAGAAATTAATTAATTAATGCGCCA	1140
Qy	1141	AGAACTCTAAGGTATATTTTGACAAATCATTCATGACAACTTTTAAAAAAGAAAGATCAGAGA	1200
Db	1141	AGAACTCTAAGGTATATTTTGACAAATCATTCATGACAACTTTTAAAAAAGAAAGATCAGAGA	1200
Qy	1201	GGCATTAAGAAAGCAGGAGGAAACAGGGAGACAGAAACACTTGTGGCCCAAGAGAACAA	1260
Db	1201	GGCATTAAGAAAGCAGGAGGAAACAGGGAGACAGAAACACTTGTGGCCCAAGAGAACAA	1260
Qy	1261	ACAAAGGCTCTTAAGCACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Db	1261	ACAAAGGCTCTTAAGCACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1320
Qy	1321	AAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAGGGATTAAGAAAG	1380
Db	1321	AAAAAAGACAGAGAGAGAGAGACAGAGACAGAGAGACAGAGAGGCGAGGGATTAAGAAAG	1380
Qy	1381	GAGAGAGGGGTGAGAGAGACACAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA	1440
Db	1381	GAGAGAGGGGTGAGAGAGACACAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA	1440
Qy	1441	GAACTCAACAGAGATGGAAGAAAGACTCTGAGAAAAAACACAGAGCAAAAGATGGAAGAGG	1500
Db	1441	GAACTCAACAGAGATGGAAGAAAGACTCTGAGAAAAAACACAGAGCAAAAGATGGAAGAGG	1500
Qy	1501	AGTATCGAGGGTGAACACAGTGGTGAATGAGCAAAATCAGAGAAAGAAAGCAAGCA	1560
Db	1501	AGTATCGAGGGTGAACACAGTGGTGAATGAGCAAAATCAGAGAAAGAAAGCAAGCA	1560
Qy	1561	TCCAGGCGCCAAAGTAAGTGAACCAAGATTGGTGAAGCCAGATCTTTAAGCTGGGGG	1620
Db	1561	TCCAGGCGCCAAAGTAAGTGAACCAAGATTGGTGAAGCCAGATCTTTAAGCTGGGGG	1620
Qy	1621	AGGCAAGGGAAGGGGCTGGCTTGGCTTCGGAAGACCTCTCCCATTTCTTCGGGGCAGGAG	1680
Db	1621	AGGCAAGGGAAGGGGCTGGCTTGGCTTCGGAAGACCTCTCCCATTTCTTCGGGGCAGGAG	1680

QY	1681	GTAGGAGTGAATTCCTCCGACCTGGGTGGGGGTGCTCTGGGGGTGGAGATATGGGGAGCA	1741
Db	1681	GTAGGAGTGAATTCCTCCGACCTGGGTGGGGGTGCTCTGGGGGTGGAGATATGGGGAGCA	1740
QY	1741	GGAGGAGCTAATTCCTAAGGCCGATATAGGCACTCAATTCGCCGGGAAATGTGCCCAAGGAG	1800
Db	1741	GGAGGAGCTAATTCCTAAGGCCCGATATAGGCACTCAATTCGCCGGGAAATGTGCCCAAGGAG	1800
QY	1801	CATGGGTGGTTATPACTCAGGCCCGGTGCCAAGGCCCAGAGAGAGCAGTGGCCACGA	1860
Db	1801	CAGTGGGTGGTTATPACTCAGGCCCGGTGCCAAGGCCCAGAGAGAGCAGTGGCCACGA	1860
QY	1861	AGGCAAGGCGCTGAGAAAGTCGGGCGTAGCTGGGAGCAATCCCCACCCCTACCTCG	1920
Db	1861	AGGCAAGGCGCTGAGAAAGTCGGGCGTAGCTGGGAGCAATCCCCACCCCTACCTCG	1920
QY	1921	GGGACAGGGCAAGTAGAGCTGTGAGAGTGCTCAGCAGGCAGAGGAAGAGAGTGTCT	1980
Db	1921	GGGACAGGGCAAGTAGAGCTGTGAGAGTGCTCAGCAGGCAGAGGAAGAGAGTGTCT	1980
QY	1981	GTGGGTCTTGACCCCAATCTTTCTGTGTCCCTCTTGGCTGTCTGGAGAGCTGTAGA	2040
Db	1981	GTGGGTCTTGACCCCAATCTTTCTGTGTCCCTCTTGGCTGTCTGGAGAGCTGTAGA	2040
QY	2041	CTCCATCTTTCTGAATCTATAGCTGGGTCTCAGCGCAGTSCCAGATGGTGGCCGTG	2100
Db	2041	CTCCATCTTTCTGAATCTATAGCTGGGTCTCAGCGCAGTSCCAGATGGTGGCCGTG	2100
QY	2101	CTGTGTGTTCTCTCTACCTGGGAAAATAAGATAGGGAGAGGAGGGAGTGGTTAAG	2160
Db	2101	CTGTGTGTCTCTCTACCTGGGAAAATAAGATAGGGAGAGGAGGGAGTGGTTAAG	2160
QY	2161	GCTCCCGGAAATCGCTGGGCTCTCCCAACCCCTGTACAATCCCAATCCAGGTGCAGCGGC	2220
Db	2161	GCTCCCGGAAATCGCTGGGCTCTCCCAACCCCTGTACAATCCCAATCCAGGTGCAGCGGC	2220
QY	2221	ATGCGTACAGACAGACCCCTCCGTGATGTGGGTCTGTGTCTGTATCAACGCTTGCTT	2280
Db	2221	ATGCGTACAGACAGACCCCTCCGTGATGTGGGTCTGTGTCTGTATCAACGCTTGCTT	2280
QY	2281	CTGGGGGTGCACAGTTAACAGAACTCTGGGTGGGAGGGTTGTGGGAATGGGAGACCTGT	2340
Db	2281	CTGGGGGTGCACAGTTAACAGAACTCTGGGTGGGAGGGTTGTGGGAATGGGAGACCTGT	2340
QY	2341	CTTGCAGGCACATAAGAGGCTGTGCCCTGGGGAATGTGTGAGCTGTGGGATGACTCCGG	2400
Db	2341	CTTGCAGGCACATAAGAGGCTGTGCCCTGGGGAATGTGTGAGCTGTGGGATGACTCCGG	2400
QY	2401	GACCGGGTGAATGTGAAGTCTGTGTCTGTACTTGTGTGTGTGCAATCGATATGTGGCCCTGT	2460
Db	2401	GACCGGGTGAATGTGAAGTCTGTGTCTGTACTTGTGTGTGTGCAATCGATATGTGGCCCTGT	2460
QY	2461	GACTGCCACGGTGTGTGTGGGAGAGGGGAAATGCTTTTCCATATACAGGTATCACTGTGCG	2520
Db	2461	GACTGCCACGGTGTGTGTGGGAGAGGGGAAATGCTTTTCCATATACAGGTATCACTGTGCG	2520
QY	2521	CAGGTGGCACTGACCTTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAATTTAAG	2580
Db	2521	CAGGTGGCACTGACCTTTTGAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGAATTTAAG	2580
QY	2581	ATTGTGTGTGTCTCAACAGCTGT	2640
Db	2581	ATTGTGTGTGTCTCAACAGCTGT	2640
QY	2641	GTGTTTGGCTGT	2700
Db	2641	GTGTTTGGCTGT	2700
QY	2701	TCCCTGAAGTCCGGGAATGGGTGCACAAAGTGTGTATCATCATTGAAAGCTGTGTACT	2760
Db	2701	TCCCTGAAGTCCGGGAATGGGTGCACAAAGTGTGTATCATCATTGAAAGCTGTGTACT	2760

QY	2761	GTGTGCTGCTTGCAGGCGCATTAATGTGATTTGTGGCTGAGTGTGAGCTTAATGGATGCCGTA	2820
Db	2761	GTGTGCTGCTTGCAGGCGCATTAATGTGATTTGTGGCTGAGTGTGAGCGTTAATGATGCCGTA	2820
QY	2821	TTTGTGACCGGTGTGACTACCTGAGCCTGTGTGTAGGGGGTGCCTGTAAATGTGACTGTGTGTG	2880
Db	2821	TTTGTGACCGGTGTGACTACCTGAGCCTGTGTGTAGGGGGTGCCTGTAAATGTGACTGTGTGTG	2880
QY	2881	TCTGTGTGAGGCGCGGTGTAAATGTCTAAGTATGTGTGATGTGTGACAGCTGTGTGTCTGGAGT	2940
Db	2881	TCTGTGTGAGGCGCGGTGTAAATGTCTAAGTATGTGTGATGTGTGACAGCTGTGTGTCTGGAGT	2940
QY	2941	TTCTGTCTCTGCTGTGAGGGGAAATAGAGGGTGCAGAGGGTACCTATCTTCTGGAGATGGGTGC	3000
Db	2941	TTCTGTCTCTGCTGTGAGGGGAAATAGAGGGTGCAGAGGGTGTAGCTATCTTCTGGAGATGGGTGC	3000
QY	3001	CAGGTGACTGACTTGCAGGTGTGTGCTGTGTGTGTCAGAAAGATATGTGGACGTCTAAACATC	3060
Db	3001	CAGGTGACTGACTTGCAGGTGTGTGCTGTGTGTGTCAGAAAGATATGTGGACGTCTAAACATC	3060
QY	3061	TGTGCACACAGGCGCATCTGTGCTGTTGAGACACTGTGATGAGGGGTGTGGATCCC	3120
Db	3061	TGTGCACACAGGCGCATCTGTGCTGTTGAGACACTGTGATGAGGGGTGTGGATCCC	3120
QY	3121	GCTAGGCTGCCCGGGAGCGTGTGTACTGTGAGACAGACAGCTGTATGTAGTGCACCTGTG	3180
Db	3121	GCTAGGCTGCCCGGGAGCGTGTGTACTGTGAGACAGACAGCTGTATGTAGTGCACCTGTG	3180
QY	3181	GAGGCAACATGAGGCGTGTCTGTGACAAACTGTGCGTGTGTGGCTGTAACTGTGACCTGTG	3240
Db	3181	GAGGCAACATGAGGCGTGTCTGTGACAAACTGTGCGTGTGTGGCTGTAACTGTGTGTGTC	3240
QY	3241	GCTGTGTCTTCTGGGGTGTGATCTGTGAATGATGTGTGTGTCAGAGGCGCATCAGCAAGGGTAA	3300
Db	3241	GCTGTGTCTTCTGGGGTGTGATCTGTGAATGATGTGTGTGTCAGAGGCGCATCAGCAAGGGTAA	3300
QY	3301	GAAACAAGGCGGGCGCGGTGCCTACCGCTGTAAATCCCAAGCTTTTGGAGAGCGGAGGCA	3360
Db	3301	GAAACAAGGCGGGCGCGGTGCCTACCGCTGTAAATCCCAAGCTTTTGGAGAGCGGAGGCA	3360
QY	3361	GGCGGATCACTGTAGGTGTGGGAGATTCAGGCGCAAGCTGTGACCAATGTGAGAAACCCCGTCT	3420
Db	3361	GGCGGATCACTGTAGGTGTGGGAGATTCAGGCGCAAGCTGTGACCAATGTGAGAAACCCCGTCT	3420
QY	3421	CTACTAAATAATACAAAAAATTAGCTGTGTGTGTGGCGGCTGCTGTAAATCCCACTATCTC	3480
Db	3421	CTACTAAATAATACAAAAAATTAGCTGTGTGTGTGGCGGCTGCTGTAAATCCCACTATCTC	3480
QY	3481	GGGAGACTGGGGCAGAAAAATCGCTGTAAACCGGGAGGTGTGAGGTTGCGGTAGCCGAGA	3540
Db	3481	GGGAGACTGGGGCAGAAAAATCGCTGTAAACCGGGAGGTGTGAGGTTGCGGTAGCCGAGA	3540
QY	3541	TGCGGCCCATTTGCATCTCCAGCTGTGGGCAACAAGCGAAATTCGCTCTCGAAGAAAAAAA	3600
Db	3541	TGCGGCCCATTTGCATCTCCAGCTGTGGGCAACAAGCGAAATTCGCTCTCGAAGAAAAAAA	3600
QY	3601	GAATAAAAAAAGGGGTAAAGAACAGTGAATTGGGCAACGGAGGACGTGATGGAATGGGGC	3660
Db	3601	GAATAAAAAAAGGGGTAAAGAACAGTGAATTGGGCAACGGAGGACGTGATGGAATGGGGC	3660
QY	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGTGAGAGAGAGATTGACAGGATTTGAGAAAGGCAT	3720
Db	3661	ATGCATGTAGTCTGTAGGTCTGTGTGTGTGAGAGAGAGATTGACAGGATTTGAGAAAGGCAT	3720
QY	3721	GTTTTTCATCTGAGAAATTCAGAAACCTTAGGCTGTCTTCCCTCCATGTGGCCCCCTAAG	3780
Db	3721	GTTTTTCATCTGAGAAATTCAGAAACCTTAGGCTGTCTTCCCTCCATGTGGCCCCCTAAG	3780
QY	3781	CTGAGGCCCTTCTTTCCTGTGCTGTGCTGTGGAAACCTTAGCTCCGCGCATAGAGCTGTAGCC	3840
Db	3781	CTGAGGCCCTTCTTTCCTGTGCTGTGCTGTGGAAACCTTAGCTCCGCGCATAGAGCTGTAGCC	3840
QY	3841	CCACCTCTTCTCTCAACACAGCCCGCTTAGGCAACTCTAGTGAACCCCGCTTAGGCA	3900

Db	3841	CCACCTCTCTTCC	TCACCA	CGCCCT	TAGGCA	GCAC	GTCTA	GTGAG	CAACCCCG	CTTAAG	CCCA	3900
Oy	3901	CACCCCTTTGGG	CCAGGCT	CA	CCCCCT	ATTC	GTGGG	AT	ACCTT	TAGA	AA	3960
Db	3901	CACCCCTTTGGG	CCAGGCT	CA	CCCCCT	ATTC	GTGGG	AT	ACCTT	TAGA	AA	3960
Oy	3961	AGTCAGAGCT	TTTTTTTTTTTTTTTTTTTT	TTGGAG	ACGT	CTTGCT	CTCTCT	CCAGG	CTGAG			4020
Db	3961	AGTCAGAGCT	TTTTTTTTTTTTTTTTTTTT	TTGGAG	ACGT	CTTGCT	CTCTCT	CCAGG	CTGAG			4020
Oy	4021	TGCAGTGGCGT	GTATCT	CGGCT	CA	CTTCG	CTCC	CAGG	TTCA	GTATTC	TCGT	4080
Db	4021	TGCAGTGGCGT	GTATCT	CGGCT	CA	CTTCG	CTCC	CAGG	TTCA	GTATTC	TCGT	4080
Oy	4081	GCCTCCAC	CTTC	TAGAG	CTGGAG	ATTA	CAGG	GGCGG	CA	CA	CGCCTGGG	4140
Db	4081	GCCTCCAC	CTTC	TAGAG	CTGGAG	ATTA	CAGG	GGCGG	CA	CA	CGCCTGGG	4140
Oy	4141	GTGTCTTTAG	TAGAG	CAGGG	TTTCA	CTTGTT	GGC	CAGG	CTGTG	CTCA	ATCCCA	4200
Db	4141	GTGTCTTTAG	TAGAG	CAGGG	TTTCA	CTTGTT	GGC	CAGG	CTGTG	CTCA	ATCCCA	4200
Oy	4201	TCAGGTAT	CCGCC	CA	CTTGCG	CTCC	CAGAG	GTGCT	GGG	TTA	CAGG	4260
Db	4201	TCAGGTAT	CCGCC	CA	CTTGCG	CTCC	CAGAG	GTGCT	GGG	TTA	CAGG	4260
Oy	4261	CCCCAGG	CCAAAG	CTGAG	CTCTTT	TPA	GGAG	ACT	CTAA	CAT	GTAA	4320
Db	4261	CCCCAGG	CCAAAG	CTGAG	CTCTTT	TPA	GGAG	ACT	CTAA	CAT	GTAA	4320
Oy	4321	CTAACTA	AGTCA	ATTCCAA	CCCCCT	CTTG	CC	CTG	CCCA	CGGCTGA	CCCA	4380
Db	4321	CTAACTA	AGTCA	ATTCCAA	CCCCCT	CTTG	CC	CTG	CCCA	CGGCTGA	CCCA	4380
Oy	4381	TGACCCCA	CTTTCTT	TAGAC	CA	GTTCAT	TCC	CTA	AA	AGCCT	GGTCTC	4440
Db	4381	TGACCCCA	CTTTCTT	TAGAC	CA	GTTCAT	TCC	CTA	AA	AGCCT	GGTCTC	4440
Oy	4441	GCTCAG	CCCCCA	CAGGCTT	GGGAC	TAC	ACCCT	CGA	AGCTT	GC	CAGGA	4500
Db	4441	GCTCAG	CCCCCA	CAGGCTT	GGGAC	TAC	ACCCT	CGA	AGCTT	GC	CAGGA	4500
Oy	4501	TTTAC	CCTCA	CATGTAG	TTCTA	CCCA	ATTC	CAGAA	TCTGT	AGGTC	CAGTTA	4560
Db	4501	TTTAC	CCTCA	CATGTAG	TTCTA	CCCA	ATTC	CAGAA	TCTGT	AGGTC	CAGTTA	4560
Oy	4561	GTAAC	CCCTA	CTGAG	CTGGG	CTCTGT	CTT	GAG	CTT	GAG	CTTG	4620
Db	4561	GTAAC	CCCTA	CTGAG	CTGGG	CTCTGT	CTT	GAG	CTT	GAG	CTTG	4620
Oy	4621	CTCTTAT	TTCTC	CA	AGGCTCG	CCCCCT	CCCT	CAG	ATTC	CA	GA	4680
Db	4621	CTCTTAT	TTCTC	CA	AGGCTCG	CCCCCT	CCCT	CAG	ATTC	CA	GA	4680
Oy	4681	GGTCTG	GCCTCTT	GAGCT	GTAA	ACCAC	CCCC	CAG	CCCGG	CTCT	GAG	4740
Db	4681	GGTCTG	GCCTCTT	GAGCT	GTAA	ACCAC	CCCC	CAG	CCCGG	CTCT	GAG	4740
Oy	4741	CAAC	CCATTTT	CCGTT	CC	CAG	CA	TGTTCT	CG	CA	CA	4800
Db	4741	CAAC	CCATTTT	CCGTT	CC	CAG	CA	TGTTCT	CG	CA	CA	4800
Oy	4801	CTCTTA	CAC	CGTGG	CCCTCT	GGGAG	CA	CA	AGGACT	CTGGG	GGGAG	4860
Db	4801	CTCTTA	CAC	CGTGG	CCCTCT	GGGAG	CA	CA	AGGACT	CTGGG	GGGAG	4860
Oy	4861	CCCGGT	CGA	TGACAG	AGCGG	CGAT	CAT	CA	TGGA	TCCG	ACTG	4920
Db	4861	CCCGGT	CGA	TGACAG	AGCGG	CGAT	CAT	CA	TGGA	TCCG	ACTG	4920
Oy	4921	AGCGGT	GGCAGG	CGCG	CTGTTG	CTA	AGG	CCCA	CCAG	CTTAC	TGCGG	4980
Db	4921	AGCGGT	GGCAGG	CGCG	CTGTTG	CTA	AGG	CCCA	CCAG	CTTAC	TGCGG	4980

Qy 721 AATATGAGTTTGGCCAGGCAACGGTGTCTCATGCTGTAAATTCACAGACTTTGGGAGGC 780
Db 721 AATATGAGTTTGGCCAGGCAACGGTGTCTCATGCTGTAAATTCACAGACTTTGGGAGGC 780
Qy 781 AGAGTGGGTGAATCATTTGAGGCCAGAGATTGTGAGACAGACTGGCCAAATGTTGAAA 840
Db 781 AGAGTGGGTGAATCATTTGAGGCCAGAGATTGTGAGACAGACTGGCCAAATGTTGAAA 840
Qy 841 CTGTGCTTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
Db 841 CTGTGCTTTAACTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 900
Qy 901 ACCGTGAATCCCAAGCTATGCTGGAGGCTGAGGCGAGAGTCACTTGAACCTTGAAGCG 960
Db 901 ACCGTGAATCCCAAGCTATGCTGGAGGCTGAGGCGAGAGTCACTTGAACCTTGAAGCG 960
Qy 961 GAGGTTGAGTGGGCCAGATCATCAACGCTCCAGCTGGGCGACAGAGCAAGACT 1020
Db 961 GAGGTTGAGTGGGCCAGATCATCAACGCTCCAGCTGGGCGACAGAGCAAGACT 1020
Qy 1021 CTGTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Db 1021 CTGTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
Qy 1081 AAAAAAATGCTGTCAACAAATTAAGCAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Db 1081 AAAAAAATGCTGTCAACAAATTAAGCAGAGTGAATTAAGAAAAATTAATGGGCCA 1140
Qy 1141 AGAATCTAAGGTTATTTTGAACAATCATTCAGAACCTTTAAAAAAGAAATCAGAGA 1200
Db 1141 AGAATCTAAGGTTATTTTGAACAATCATTCAGAACCTTTAAAAAAGAAATCAGAGA 1200
Qy 1201 GGCATAGAAAAGACAGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
Db 1201 GGCATAGAAAAGACAGGAGGAAACAGGAGACAGAAACACTGTGGCCCAAGAGAACAA 1260
Qy 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 ACAAGGCTCTTAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Qy 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 AAAAAACAGAGAGAGAGACAGAGACAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAG 1380
Qy 1381 GAGAGAGGAGTGAAGAGACAGAGATTTGAGAGAGACTCAGAAAGTATGCGAGGGA 1440
Db 1381 GAGAGAGGAGTGAAGAGACAGAGATTTGAGAGAGACTCAGAAAGTATGCGAGGGA 1440
Qy 1441 GAACCAACAGAGAGATGGAAGAACTCTGAGAAAAAACAGAGACAAAGATGGAAGAG 1500
Db 1441 GAACCAACAGAGAGATGGAAGAACTCTGAGAAAAAACAGAGACAAAGATGGAAGAG 1500
Qy 1501 AGTATCGAGGAGTGAACAGACAGTGTGGAATGAGCAAAATTCAGAGAAAGAAAGCA 1560
Db 1501 AGTATCGAGGAGTGAACAGACAGTGTGGAATGAGCAAAATTCAGAGAAAGAAAGCA 1560
Qy 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTGGAAGCCAGATCCTTAAAGGCTGGGG 1620
Db 1561 TCCAGGCGCCAAAGATAGTGAACCAAGATTGTGGAAGCCAGATCCTTAAAGGCTGGGG 1620
Qy 1621 AGGCAAGGAGAGGGGCTGGCTGGCTCTCGAGAGACCCCTCCCATTTCTCCGGGCGAG 1680
Db 1621 AGGCAAGGAGAGGGGCTGGCTGGCTCTCGAGAGACCCCTCCCATTTCTCCGGGCGAG 1680
Qy 1681 GTAGGAGATGACATTCGGAATGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
Db 1681 GTAGGAGATGACATTCGGAATGGGTGGGGGTGCTCTGGGGGTGAGATAGGGGAGCA 1740
Qy 1741 GGAGAGCTATTTGCTAAGGCCGATAGGCACTTATGCTCCGGGAGATGTCGCCAGGAG 1800
Db 1741 GGAGAGCTATTTGCTAAGGCCGATAGGCACTTATGCTCCGGGAGATGTCGCCAGGAG 1800

Qy 1801 CAGTGGGTGTTATTAATCAGGCCCGGTGCCAGAGCCAGAGAGGAGAGGAGTGGCCAGGA 1860
Db 1801 CAGTGGGTGTTATTAATCAGGCCCGGTGCCAGAGCCAGAGAGGAGAGGAGTGGCCAGGA 1860
Qy 1861 AGGCAAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCACTTCACTGG 1920
Db 1861 AGGCAAGGCTGAGAAAGTCTGGGCTGAGCTGGAGCAAAATCCCACTTCACTGG 1920
Qy 1921 GGGACAGGAGAGTGAAGCTTGTGAGAGGCTGTGAGAGAGAGAGAGAGAGAGAGTGT 1980
Db 1921 GGGACAGGAGAGTGAAGCTTGTGAGAGGCTGTGAGAGAGAGAGAGAGAGAGAGTGT 1980
Qy 1981 GTGGGCTGTGACCCAGATCTTCTGTGCTCCCTCTGCTGCTGTGAGAGCTGTAGA 2040
Db 1981 GTGGGCTGTGACCCAGATCTTCTGTGCTCCCTCTGCTGCTGTGAGAGCTGTAGA 2040
Qy 2041 CTCTATCTTCTGAATCTATAGTGTGCTGGGTCTCAGGCGAGTGCAGATGCTGGCCCTG 2100
Db 2041 CTCTATCTTCTGAATCTATAGTGTGCTGGGTCTCAGGCGAGTGCAGATGCTGGCCCTG 2100
Qy 2101 CTGTGTGTTCTCTCTTACCTGGGAAATTAAGTAAAGGAGAGAGAGAGAGAGTAAAG 2160
Db 2101 CTGTGTGTTCTCTCTTACCTGGGAAATTAAGTAAAGGAGAGAGAGAGAGAGTAAAG 2160
Qy 2161 GCTCCCGGAGTGGCTGGGCTCCAACTCTGACATTCCTCCATCCAGGTGACGCGCC 2220
Db 2161 GCTCCCGGAGTGGCTGGGCTCCAACTCTGACATTCCTCCATCCAGGTGACGCGCC 2220
Qy 2221 ATGGCTACAGCAAGACCCCTGGATGAGGAGTGTGCTGTGATCAGAGCTTGTCTT 2280
Db 2221 ATGGCTACAGCAAGACCCCTGGATGAGGAGTGTGCTGTGATCAGAGCTTGTCTT 2280
Qy 2281 CTGGGGGTGACAGGTAAACAGAACTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGTGT 2340
Db 2281 CTGGGGGTGACAGGTAAACAGAACTCTGGGAGAGAGAGAGAGAGAGAGAGAGAGTGT 2340
Qy 2341 CTGTGCGGCACTAGAGGCTGTCCCTGTGGGAACTGTGTGAGGCTGGGATGACTCCGG 2400
Db 2341 CTGTGCGGCACTAGAGGCTGTCCCTGTGGGAACTGTGTGAGGCTGGGATGACTCCGG 2400
Qy 2401 GACCGGAGTGAATGAGTCTGTGTGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Db 2401 GACCGGAGTGAATGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2460
Qy 2461 GACTGCCACGAGT 2520
Db 2461 GACTGCCACGAGT 2520
Qy 2521 CAGGTGACCTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Db 2521 CAGGTGACCTGACCTTTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2580
Qy 2581 ATTGTGTGTGCTTCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
Db 2581 ATTGTGTGTGCTTCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
Qy 2641 GT 2700
Db 2641 GT 2700
Qy 2701 TCCCTGAGGTCCTGGGAGTTGCTGACCAAAAGGTGATCACCATGGAAGAGCTGTGACT 2760
Db 2701 TCCCTGAGGTCCTGGGAGTTGCTGACCAAAAGGTGATCACCATGGAAGAGCTGTGACT 2760
Qy 2761 GTGTGCTGTGTGACGAGGATTAATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2820
Db 2761 GTGTGCTGTGTGACGAGGATTAATGTGATTTGTGTGTGTGTGTGTGTGTGTGTGTGT 2820
Qy 2821 TTTGTGTGACGCTGTGACCTTGAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Db 2821 TTTGTGTGACGCTGTGACCTTGAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2880
Qy 2881 TCTGTGTGAGGCGGTGAATGTCTATGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940

Db 2881 TCTGTGTGAGGCGCGTGAATGCTACTGATGTGTGATGTGTGACAGCTGTGTCTGTGAGT 2940
QY 2941 TTCTGTCTGTGCGTGGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGTGGTGC 3000
Db 2941 TTCTGTCTGTGCGTGGAGGGATAGAGGGTGCAGGGGTAGCTATCTCTGGAGATGTGGTGC 3000
QY 3001 CAGGTGACTGACTTGTGAGTGTGTGCTGTGTGCAAGAGATATGTGGCACTGTGAACATC 3060
Db 3001 CAGGTGACTGACTTGTGAGTGTGTGCTGTGTGCAAGAGATATGTGGCACTGTGAACATC 3060
QY 3061 TGTGTGACACAGGCACTGTGTGCTGTGTGCAAGAGATATGTGGCACTGTGAACATC 3120
Db 3061 TGTGTGACACAGGCACTGTGTGCTGTGTGCAAGAGATATGTGGCACTGTGAACATC 3120
QY 3121 GCTAGGCTGCGCGGAGCGGTGTGATCCTGTGAGACAGAGCTGTATGTAGCTGACCTGTG 3180
Db 3121 GCTAGGCTGCGCGGAGCGGTGTGATCCTGTGAGACAGAGCTGTATGTAGCTGACCTGTG 3180
QY 3181 GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3240
Db 3181 GAGGCAACATGTGGCGTGTGTGCAAGAACTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3240
QY 3241 GCGTGTGTCTTGTGGGTGTGAGTGTGTGTATGTATGTGTGTGCTGTGTGCTGTGTGCTG 3300
Db 3241 GCGTGTGTCTTGTGGGTGTGAGTGTGTGTATGTATGTGTGTGCTGTGTGCTGTGTGCTG 3300
QY 3301 GAACAGGCGCGGAGCGGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3360
Db 3301 GAACAGGCGCGGAGCGGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3360
QY 3361 GCGGATCACTGTGAGTGTGTGAGATGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3420
Db 3361 GCGGATCACTGTGAGTGTGTGAGATGTGTGCTGTGTGCTGTGTGCTGTGTGCTGTGTGCTG 3420
QY 3421 CTACTTAAATAATTAATAATAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3480
Db 3421 CTACTTAAATAATTAATAATAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3480
QY 3481 GGGAGACTGGGGCAGAAAAATGCTTGAACCGGAGAGGTGTGTGTGTGTGTGTGTGTGTG 3540
Db 3481 GGGAGACTGGGGCAGAAAAATGCTTGAACCGGAGAGGTGTGTGTGTGTGTGTGTGTGTG 3540
QY 3541 TCGCGCATTTGCACTCAGCTGTGGGCAACAAGAGCAAACTCCGTCTCGAAAGAAAAA 3600
Db 3541 TCGCGCATTTGCACTCAGCTGTGGGCAACAAGAGCAAACTCCGTCTCGAAAGAAAAA 3600
QY 3601 GAAAAAAGGATTAAGAACAGTG 3660
Db 3601 GAAAAAAGGATTAAGAACAGTG 3660
QY 3661 ATGCATG 3720
Db 3661 ATGCATG 3720
QY 3721 GTTTTCATCTGAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 3780
Db 3721 GTTTTCATCTGAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAATTCAG 3780
QY 3781 CTGAGGCTTCTTCTGAGTG 3840
Db 3781 CTGAGGCTTCTTCTGAGTG 3840
QY 3841 CCACCTCTTCTCTCAACACGCGCTGTGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3900
Db 3841 CCACCTCTTCTCTCAACACGCGCTGTGAGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3900
QY 3901 CACCCCTTTGTGGGCAAGGCTCAACCCCTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3960
Db 3901 CACCCCTTTGTGGGCAAGGCTCAACCCCTTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 3960
QY 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGTGAGACAGTGTGTGTGTGTGTGTGTGTGTGTG 4020

Db 3961 AGTCAGAGCTTTTTTTTTTTTTTTTTTTTGTGAGACAGTGTGTGTGTGTGTGTGTGTGTGTG 4020
QY 4021 TGCAGTGGGATATCTCGGCTCAGTCAACCTGTGCTCCAGGTTCAAGTATTTCTGT 4080
Db 4021 TGCAGTGGGATATCTCGGCTCAGTCAACCTGTGCTCCAGGTTCAAGTATTTCTGT 4080
QY 4081 GCGTCACTCTGTAGTG 4140
Db 4081 GCGTCACTCTGTAGTG 4140
QY 4141 GTGTCTTTG 4200
Db 4141 GTGTCTTTG 4200
QY 4201 TCAGGTGATCGGCGCACTG 4260
Db 4201 TCAGGTGATCGGCGCACTG 4260
QY 4261 CCCAGCCCAAGTG 4320
Db 4261 CCCAGCCCAAGTG 4320
QY 4321 CTACTTG 4380
Db 4321 CTACTTG 4380
QY 4381 TGACCCCACTTG 4440
Db 4381 TGACCCCACTTG 4440
QY 4441 GCTCAGCCCGCAAGCTTTG 4500
Db 4441 GCTCAGCCCGCAAGCTTTG 4500
QY 4501 TTTTACCTTG 4560
Db 4501 TTTTACCTTG 4560
QY 4561 GTTAACTTG 4620
Db 4561 GTTAACTTG 4620
QY 4621 CTCTTATTTCTGAGGCGCTG 4680
Db 4621 CTCTTATTTCTGAGGCGCTG 4680
QY 4681 GGTCTG 4740
Db 4681 GGTCTG 4740
QY 4741 CAACCAATTTTCTGTTCCAGAGCAATG 4800
Db 4741 CAACCAATTTTCTGTTCCAGAGCAATG 4800
QY 4801 CCTTAAACCTG 4860
Db 4801 CCTTAAACCTG 4860
QY 4861 CCCGCTG 4920
Db 4861 CCCGCTG 4920
QY 4921 AGCGGTG 4980
Db 4921 AGCGGTG 4980
QY 4981 TGCATTCACAGTG 5000
Db 4981 TGCATTCACAGTG 5000

RESULT 5
ACH80345

ID ACH80345 standard; DNA; 525 BP.
 AC ACH80345;
 DT 29-JUL-2004 (first entry)
 DE Human genome derived single exon probe #13540.
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KM alternative splicing event; genomic alteration.
 XX Homo sapiens.
 OS
 PN US2003194704-A1.
 PD 16-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00029386.
 XX
 PR 03-APR-2002; 2002US-00029386.
 PA (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 PI Penn SG, Rank DR, Hanzel DK;
 DR WPI; 2004-119264/12.
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 PS Claim 15; SEQ ID NO 13540; 80pp: English.
 XX
 CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridises under high stringency conditions to a nucleic acid molecule
 CC addressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterising
 CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/Sequence.html?DocID=20030194704
 XX

	Query Match	9.4%; Score 472; DB 12; Length 525;
	Best Local Similarity	100.0%; P-Id. No. 4,3e-76;
	Matches 472; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	4529	TCCAGGAATCTGTAGAGTCCAGTTAAGTCCAGTAACCTTACCCTGAGCCTGGCTCTGTG
Dd	1	TCGAGGAATCTGTAGAGTCCAGTTAAGTCCAGTAACCTTACCCTGAGCCTGGCTCTGTG
Oy	4589	CTTGAGCTTGAAGCTGGGCTTTGAGAAGGTGCCACTCTTATTCTCCAGGCCCTTGCCCCCGC
Dd	61	CTTGAGCTTGAAGCTGGGCTTTGAGAAGGTGCCACTCTTATTCTCCAGGCCCTTGCCCCCGC
Oy	4669	CCCTCAGATCTGACAGAACCAACCCCTTAGCTGGTCTGGCTCTTGAGTCTGAAGACCCAC
Dd	121	CCCTCAGATCTGACAGAACCAACCCCTTAGCTGGTCTGGCTCTTGAGTCTGAAGACCCAC
Oy	4709	CCCCAGCCCAAGCCCCTCTGAGCCCGGCCCAACCATTTTCCGTTCCAGAGCATGT
Dd	181	CCCCAGCCCAAGCCCCTCTGAGCCCGGCCCAACCATTTTCCGTTCCAGAGCATGT
Oy	4769	TCTGGCCAACAATGATGTTTTCTGTGACCAACCCCTCTPAACAACGCTGCCTCTGGAGCAA
Dd	241	TCTGGCCAACAATGATGTTTTCTGTGACCAACCCCTCTPAACAACGCTGCCTCTGGAGCAA
Oy	4829	CCAGGACCTGGGAGCTGGGGCGGGGGAAGAACCGCCGGTCCGATACAGACAGACCCGAT
Dd	301	CCAGGACCTGGGAGCTGGGGCGGGGGAAGAACCGCCGGTCCGATACAGACAGACCCGAT
Oy	4889	CATCAATGATCCGACCTGCGATATGACACCCACGCTGGCAGGCGCGCTGTGTCTAAG
Dd	361	CATCAATGATCCGACCTGCGATATGACACCCACGCTGGCAGGCGCGCTGTGTCTAAG
Oy	4949	GCCCAACCAAGCTCTACTGCGGGGCGGTGTGTGTCATCCACAGTGGCTGCTC 5000
Dd	421	GCCCAACCAAGCTCTACTGCGGGGCGGTGTGTGTCATCCACAGTGGCTGCTC 472
RESULT 6		
ID	ACN44942	standard; DNA; 41454 BP.
AC	ACN44942;	
DT	18-NOV-2004	(first entry)
DE	Human genomic sequence hCG21040.	
KW	Cyclostatic; carcinoma; lymphoma; cancer; human; gene; sb.	
OS	Homo sapiens.	
PX	WO2003073826-A2.	
PN	12-SEP-2003.	
PD	28-FEB-2003; 2003WO-US006235.	
PR	01-MAR-2002; 2002US-00087192.	
PA	(SAGR-) SAGRES DISCOVERY.	
XX	Morris DW;	
PI	WI; 2003-328604/31.	
DR	Recombinant nucleic acid useful for diagnosis and treatment of carcinoma	
PT	comprises a nucleotide sequence.	
XX	Claim 1; SEQ ID NO 1642; Opp; English.	
XX	The present invention relates to novel DNA and protein sequences which	
CC	are associated with carcinomas. The sequences are useful for: (1) for	

OY	314	CACAGTTTACGGGTGGTCAAGATCACTCTTTTTTTTTTTTTTTTTTTTTT	373
Db	20797	CACAGTGCCTATTTTCCCTTCCAGATATGATTCCTTTTTTTTTTTTTTTTTT	20738
OY	374	TTGAGACGGAGCTTCGTCTGTCTGTCCGACAGGCTGGAGTGCAGTGGCGAGATCTCGGCTCAC	433
Db	20737	TTGAGACGGAGCTTCGTCTGTCTGTCCGACAGGCTGGAGTGCAGTGGCGAGATCTTGGCTCAC	20678
OY	434	TGCAAGCTCCGCTCCCGGGTTACAGGCAATCTTCGCTCAGCTTCCCAATGATGCTGGG	493
Db	20677	TGCAAGCTCCGCTCCCGGGTTACAGGCAATCTTCGCTCAGCTTCCCAATGATGCTGGG	20618
OY	494	ACTACAGGCGCCGCGCACTACGCGCCGCTAATTTTTTGTATTTTATAGTAGACGGGGT	553
Db	20617	ACGACAGGCGCCGCGCACTACGCGCCGCTAATTTTTTGTATTTTATAGTAGACGGGGT	20559
OY	554	TTCAACGTTTTTACCGGGATGGCCTCGATCTCTGACCTTCGTATCGCCCGCTCGGCC	613
Db	20558	TTCAACGTTTTTACCGGGATGGTCTGATCTCTGACCTTCGTATCGCCCGCACTTCGGCC	20499
OY	614	TCCCAAGTGTGGGATTTACAGGGGTGAGCCACCGGGCCCGGCATGATCATCT-----	667
Db	20498	TCCCAAGTGTGGGATTTACAGGGGTGAGCCACCGGGCCCGGCATGATTTCTTAA	20439
OY	668	-----TCTGACTAGCTGATGATGCAAGTAC	694
Db	20438	TGTGCTATTTTGGTTAACTTCTTATTTTGAAGTTTTTGGATATACATGAAACAAGCAC	20379
OY	695	CTAAGGCCATCA-GACTCTACCTTTTAAATATGCAAGTTGGGCGAGCACCGTGGCTCAT	753
Db	20378	AAATAAATATATCATCTTATCCACCGCTTAATATACAAATGTTATAGCTCAGGACGGTGACTCAC	20319
OY	754	GCTGTATATTCAGACATTTTGGAGGAGCAGAGGTGGTGAATCATCTTGAGCCAGAGATT	813
Db	20318	ACCTGTATATCCCAACATTTTGAAGGGCCGAGATGGGCACTCAC--GAGGTCAAGAGTTTC	20261
OY	814	GAGACCAAGCTGGGCAACATGTGTGAATCTGTCTTTACTTAAAAAAAAAAAAAAA	873
Db	20260	GAGACCAAGCTGGGCAACATGTGTGAATCTGTCTTTACTTAAAAAAAAAAAAAAA	20213
OY	874	AAAAATCAGCCGGGTGTCTGTGGGACACCTGTATCTCCAGCTATGCTGAAGGCTGAAGC	933
Db	20212	AAAAATATAGCTGGGTGTGTGTGGCCAGGCTGTATCTCCAGCTATGCTGAAGGCTGAAGC	20153
OY	934	ACGAGATCTCACTTGAACCTCTGAGGCGGAGGTTCAGTGGGCGGAGATCATCACGCC	993
Db	20152	ATGGAATTTCTTGAGGCCCGGAGGCGAGAGTGTGCAGTGAGTGAATCATCATTTGA	20093
OY	994	CTTCAGGCTGGGCGACAGACAAAGACTCTGTCTCAATATAATTAATTAACAAACGACAA	1053
Db	20092	GTCACAGCTGGGCGACAGACGAGACTCTCTTCTCAATATAATTAATTAATTAATTAAGTT	20033
OY	1054	GCAGTTTGTGTACTTATGTTATATCTTAAAAAAAAT	1091
Db	20032	GTTGATCTTTCACCATATCATTTAATGATGATTAAGAT	19995
RESULT 8			
ADN94799			
ID ADN94799 standard; DNA; 91352 BP.			
XX ADN94799;			
XX			
XX 12-AUG-2004 (first entry)			
XX			
XX DNA encoding human nidogen.			
XX			
XX gene; human; Antisense Therapy; Gene Therapy; nidogen;			
XX Chedick-Higashi syndrome; de.			
XX			
XX Homo sapiens			

PN	US2004097451-A1.
XX	
PD	20-MAY-2004.
XX	
PF	19-NOV-2002; 2002US-00300611.
XX	
PR	19-NOV-2002; 2002US-00300611.
XX	
PA	(ISIS-) ISIS PHARM INC.
XX	
P1	Chiang M, Dobie KM;
XX	
DR	WPI; 2004-389192/36.
XX	
PT	New compounds, particularly oligonucleotides targeted to a nucleic acid encoding nldogen, useful for treating diseases associated with nldogen, e.g. Chediak-Higashi syndrome.
CC	
PS	Claim 1; SEQ ID NO 4; 91pp; English.
XX	
CC	The invention relates to antisense oligonucleotides which are targeted to, and inhibit the expression of, a nucleic acid molecule encoding nldogen. The antisense oligonucleotides are useful for treating a disease or condition associated with nldogen, such as Chediak-Higashi syndrome. They are also useful in research and diagnostics for modulating the expression of nldogen. The present sequence represents the gene encoding human nldogen.
CC	
SQ	Sequence 91352 BP; 23543 A; 20675 C; 20522 G; 26612 T; 0 U; 0 Other;
Query Match	8.4%; Score 419; DB 12; Length 91352;
Best Local Similarity	79.4%; Pred. No. 3e-66;
Matches 552; Conservative	0; Mismatches 125; Indels 18; Gaps 4;
OY	350 TTTT TTTTTTTTTTTTTTTTTTTTTTTGAGACGGAGTCGTCTGTCGCCAAGCGTGAG 409
DB	11334 TTATTTATTATTATTATTATTATTATTGAGTAGAGTCCTCTGTCGCCAAGTTGAG 11393
OY	410 TGCAGTGC CGGAGTCTCGCTCAGCAAGCTCCGCTCCCGGTTCAAGCCATTCTCT 469
DB	11394 TGCAAGTGTGTGCTTTGGCTTAATGCAAGCTCACCTCCGGGTTCCGGCATTTCTCT 11453
OY	470 GCCTCAGCTCCCAAGTAGCTGGACTACAGGGCCCCGACCTAGCCCGGCTAATTTT 529
DB	11454 GCCTCAGCTCCCGAGTAGCTGGACTACAGAACCCGACACAGCCCGGCTAA-TTTT 11512
OY	530 TTGTATTTTTTAAAGNAGAGCGGGTTTCAACGTTTAAAGCGGGATGGCTCGATCTCTGA 589
DB	11513 TTGTATTTTTTAAAGNAGAGCGGGTTTCAACGTTTAAAGCGGGATGGCTCGATCTCTGA 11572
OY	590 CCTCGTAGATCCGCCGCCCTCGGCTCCCAAAGTGTGGATTTAACAAGCGTGAACACC 649
DB	11573 CTCTCGTAGTCAACCACTCGGCTCCCAAAGTGTGGATTTAACAAGCGTGAACACC 11632
OY	650 GCCCGGCATGATCATCTTCTTGACTATGCT---TGATGTGA CAAGTACTTAAGCCATC 705
DB	11633 ACCGGCCCCGGAATTTATTTTAAAGCAAAATCCAAATAATTTATCGTTTGTAAATGTTT 11692
OY	706 AGACTCTACCTTTAAATATG-CAGTTTGGCGCAGCACCGGTGCTCATGCTGTAAATTC 764
DB	11693 AGATGTGCTCAGAAAAGAAAGACTCTTTTGGCGGAGCAAGTGGCTCATACCTGTAAATTC 11752
OY	765 CACACATTTGGGAGGAGAGTGGGTGATTCATTGAGGCGAGGATTTGAACAGAGCT 824
DB	11753 CAGCATTTTGGAGGCTGAGGCGGCGGAGATAGCTGACGAGGTTCCAGACTTAGCTT 11812
OY	825 GGCCAACATGTGAAACTCTGTCTTTACTTAAAAA AAAAAAAAAAAAAAAAAAATCAGCC 884
DB	11813 GGCCAACATGTGAAAACCTGTGTCACT-----AAAAATATCAAAATTAAGC 11860
OY	885 GGGTGTCTGTGGGCAACCTGTAAATCCAGCTATGCTGAGAGCTGAGGCAAGAGTCAAC 944
DB	11861 AGATGTGTGTGGCAACACCTGTAAATCCAGCTATGCTGAGAGGCTGAGGCAAGAGTATTC 11920

Qy	945	TTGAACCTGGAGCGGAGTTGACAGTGGGCGCAGATCAATACCGCCCTCCAGCTGG	1004
Db	11921	TTGAATTCAGAGAGTAAAGGTTTGACATGAGCCGACGTGCACATGCACTGCAGCTGG	11986
Qy	1005	GCAGCAGAGCAAGACTGTGTCTCAATTAATAAT	1039
Db	11981	GTGACAGAGCAGACTGTGTCTCAAAAAAAAAAAT	12015
RESULT 9			
ID	ADR52890	standard; DNA; 115935 BP.	
AC	ADR52890;		
DT	18-NOV-2004	(first entry)	
DE	Drug therapy altered expressed gene #241.		
XX	drug activity monitoring; expression profile; gene expression;		
KW	peripheral blood sample; peripheral blood mononuclear cell; drug therapy;		
KW	CCI-779; immunosuppressant; rapamycin; mammalian target of rapamycin;		
XX	mTOR; ds.		
XX	Homo sapiens.		
XX	WO2004072265-A2.		
PD	26-AUG-2004.		
XX	11-FEB-2004; 2004WO-US004118.		
PF	11-FEB-2003; 2003US-0446133P.		
PR	03-APR-2003; 2003US-0459782P.		
PR	23-JAN-2004; 2004US-0538246P.		
XX	(AMMP) WYETH.		
PA	(BURC/) BURCZYNSKI M.		
PA	(TWIN/) TWINE N.		
PA	(DORN/) DORNER A J.		
PA	(TREP/) TREPICCHIO W L.		
XX	Burczynski M, Twine N, Dorner AJ, Trepicchio WL;		
PI	WPI; 2004-642301/62.		
DR			
XX			
PT	Monitoring drug activities in vivo comprises comparing an expression		
PT	profile of a gene in a peripheral blood sample of a patient before and		
PT	after drug therapy.		
XX			
PS	Disclosure; SEQ ID NO 241; 136pp; English.		
XX			
CC	The invention relates to a method of monitoring drug activities in vivo		
CC	by comparing an expression profile of at least one gene in a peripheral		
CC	blood sample of a patient to a reference expression profile of the at		
CC	least one gene, where the at least one gene is differentially expressed		
CC	in peripheral blood mononuclear cells (PBMCs) of patients who have a non-		
CC	blood disease and are subjected to a drug therapy as compared to PBMCs		
CC	isolated from the patient before the drug therapy, and where the patient		
CC	has the non-blood disease and is being treated by the drug therapy. The		
CC	method, kit, and nucleic acid array are useful for monitoring drug		
CC	activities in vivo. The drug is especially CCI-779, an ester analogue of		
CC	the immunosuppressant rapamycin which is a potent inhibitor of the		
CC	mammalian target of rapamycin (mTOR). This sequence represents a gene		
CC	expressed in PBMC altered by the drug therapy. (Note: this sequence does		
CC	not form part of the printed specification but was obtained in electronic		
CC	format from WIPO at ftp.wipo.int/pub/published_pot_sequences/).		
XX			
SQ	Sequence 115935 BP; 29948 A; 27629 C; 27920 G; 30438 T; 0 U; 0 Other;		
Query Match	8.4%; Score 418.4; DB 13; Length 115935;		
Best Local Similarity	70.2%; Pred. No. 3.9e-66;		

[illegible]

OY	340	TCACTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAAGCGAAGTCGTGCTGGCC	399
Dd	31712	TCATATTATTAATAATTTTTTTTTTTTTTTTTTTTGAAAGCCGAAGTCGTCTGTACCC	317711
OY	400	CAGCCTGAGAGCAGTAGCGAGGATCTCGGCTCATCTGCAAGCTCCGCTCCGGATTACG	459
Dd	31772	CAGGCTGGAAGTGAGTGGCGGAGACTTGGCTCACTGAAATCTCACCTCCGGATTACG	318311
OY	460	CCATTCTCTGCTCAGCTTCCCAGAATAGTGGAGCTACAGGCGCCGCCACTACGC	519
Dd	31832	CCATTCTCTGCTCAGCTTCCCAGAATAGTGGAGCTACAGGCGCCGCCACACCCG	318911
OY	520	GCTAATTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCAACGGTTTACCGGAGTGGCC	579
Dd	31892	GCTAATTTTTTTTGTATTTTTAGTAGAGACGGGGTTTCAACGGTTTACCAGATGGCTC	319511
OY	580	GATCTCTGACCTCGATCCGCGCCGCTCGGCTCCCAAAGTCTGGAGTTACAGGCGT	639
Dd	31952	GATATCTGATCTTGTGTATCGCGCTGCTGAGATCCCAAAGTCTGGAGTTACAGGCGT	320111
OY	640	GAGCCACCGCGCCCGGCGCATGATCATTTCTTGACTATGCTGATGTGACAATCTTAA	699
Dd	32012	GAGCCACCGCGCATCCGCGCGCTGTTCTGTTTCTATATTTTAT-TCAAAATTAATTVAG	320701
OY	700	GCCATCGACCTCTAACCTTTAAAATATCAAGTTTGGCGACGCGTGCCTCATGCTGT	759
Dd	32071	GCAATCA-----GCTAGGCTTGGCTCATGCTGTGCTC	321021
OY	760	AATTCAGACATTTTGGAGGCGAGAGTGGTGAATCATTTGAGGCGAGAGTTTGAGACC	819
Dd	32103	AATCCAGAC- TTGGAGGCGCAAGGAGTGAATCATTTGAGAGTTTGAGATC	321611
OY	820	AGCTGCGCAATGATGTAATCTGTCTTTACTAAAAA-----A	874
Dd	32162	AGCTGCGCAACGATGTAATCTGTCTTTACTAAAAA-----A	322211
OY	875	AAATTCAGCGCGGATGCTGGAGGCGACCTGATCCAGCTTGTGAGGCTGAGGCA	934
Dd	32222	AAATTCAGCGCGGATGCTGGAGGCGACCTGATCCAGCTTGTGAGGCTGAGGCA	322811
OY	935	CGABAGTCACTTGAACCTTGAAGCGGAGGTTGCAATGAGGCGAGATCATCACCGCC	994
Dd	32282	GGAATGATGCTTGAACCTTGAAGCGGAGGTTGCAATGAGGCGAGATCATCACCGCTAC	323411
OY	995	TCCAAGCTCGGCGCAGACAGCAAGCTGTCTCAATTAATAATAACGAACGACG	1054
Dd	32342	TCCAAGCTCGGCGCAGACAGCAAGCTGTCTCAATTAATAATAAAAAAAAAAAG	324011
OY	1055	CAGTTGTGTACCTTAGTTATATCTTAAA	1084
Dd	32402	AAGGCGGTGTTGGTGGCTCAGCTCTTAA	324311
<hr/>			
RESULT 12			
ADFe9167			
ID	ADFe9167	standard; cDNA; 216215 BP.	
XX	ADFe9167;		
DT	12-FEB-2004	(first entry)	
DE	Human MP53 nucleotide sequence SRQ ID NO:25.		
XX	p53 pathway modulating agent; MP53; p53 modulator; cytosolic;		
KM	gene therapy; cancer; human; gene; ss.		
XX	Homo sapiens.		
OS	MO2003083047-A2.		
PN	09-OCT-2003.		
PD			

PF	28-FEB-2003; 2003MO-US006025.
PR	01-MAR-2002; 2002US-0361196P.
XX	(EXEL-) EXELIXIS INC.
PA	Belyin M, Francis-Lang H, Friedman L, Ploewman GD, Heuer TS, Li D;
XX	Punkte RP;
PI	WPI; 2003-812540/76.
DR	P-PsDB; ADF69111.
XX	
PT	Identifying a candidate p53 pathway modulating agent for treating e.g.,
PT	cancer by contacting an assay system comprising a MP53 polypeptide or
PT	nucleic acid with a test agent and detecting a test agent-biased
PT	activity.
PS	
XX	
PS	Example; SEQ ID NO 25; 406pp; English.
CC	The present invention describes a method for identifying a candidate p53
CC	pathway modulating agent, which comprises: (a) providing an assay system
CC	comprising a MP53 (modulator of p53) polypeptide or nucleic acid or its
CC	fragment or derivative; (b) contacting the assay system with a test agent
CC	under conditions where the system provides a reference activity except in
CC	the presence of the test agent; and (c) detecting a test agent-biased
CC	activity, where a difference between the test agent-biased activity and
CC	the reference activity identifies the test agent as a candidate p53
CC	pathway modulating agent. Also described: (1) modulating the p53 pathway
CC	of a cell; (2) modulating the p53 pathway in a mammalian cell; and (3)
CC	diagnosing a disease in a patient. MP53 has cytostatic activity, and can
CC	be used in gene therapy. The method is useful for identifying a candidate
CC	p53 pathway modulating agent for preparing a composition for diagnosing
CC	or treating e.g., cancer. The present sequence encodes a human MP53
CC	protein, which is used in the exemplification of the present invention.
XX	
XX	Sequence 216215 BP; 64591 A; 41616 C; 43966 G; 66042 T; 0 U; 0 Other;
Query Match	8.3%; Score 417.4; DB 10; Length 216215;
Best Local Similarity	77.8%; Pred. No. 6, 4e-66;
Matches 544; Conservative	0; Mismatches 146; Indels 9; Gaps 3;
QY	347 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGCGAGCTGCTGTGCGCCAGGCTG 406
DB	117634 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGCGGGCTTGCTGTGCACCCAGGCTG 117693
QY	407 GAGTCGATGGGGGAGATCTCGGCTCACTGCGAAGCTCCGCTCCCGGTTCAAGCATTCT 466
DB	117694 GAGTCGGGTGGGTATCTCACTCACTGCAACTCTTCACTCTTAGTTCAGAAGCTATTTCT 117753
QY	467 CCTGCTCAGCCTCCCAAGTAGCTGGAGACTACAGCGGCCGACACTACGCGCGCATTT 526
DB	117754 TCTGCTCAGCCTCCCGAGTAGCTGGAGACTACAGCGCTGACACCCTCATGCCAGCTTA-- 117811
QY	527 TTTTGTATTTTAAAGAGAGCGGGGTTTACCGTTTTAGCCGGGATGGCTCTGATCTCC 586
DB	117812 TTTTGTATTTTAAAGAGAGGGGTTTACCATATCTGCGCAGGGGTGCTTTGAATCTCC 117871
QY	587 TGACCTCGTAAAGCCCGCCCTCGGCTCCCAAGTCTGGAGATTACAGCGCGAGCCAC 646
DB	117872 TGACCTTGAGATCTCTGCTCTCAAGCTCCCAAAGTCTGGAGATTACAGCGCGAGCCAC 117931
QY	647 CGCGCCCGGCATGATCATCTTCTTGACTATCTGATGGAAGTAAAGTCATCA 706
DB	117932 TGTGACTGGCCCTCATCCACTTCTCTACAACCAACCTTTTAG-----GTGTGTTTA 117985
QY	707 GACTCTACCTTTAAATAATGAGTTTGGCGCAGGACCGGTGCTCATGCTGTATTTCCA 766
DB	117986 TTTCCTCATGTAAAGAAAATCTTTCCGCGCAGGACGGGTGCTCACGCTGTATTTCCA 118045
QY	767 GCACCTTGGAGAGAGAGTGGGTGAATCATCTTAGAGCGCAGAGATTGAGACAGGCTGG 826
DB	118046 GCACCTTGGAGAGAGTGGGTGAATCATCTTAGTGAAGATTGAGACAGGCTGG 118105

QY	827	CCAAATGATGAAACCTGCTCTTACTTAAATCAAAAAAAAAAAAAAAAAATCAGCCG	886
Db	118106	CCCACTGCGAAATTCCTGCTCTTACTTAAATCAAAAAAAAAAAAAAAAAAT-AGCTGG	1181644
QY	887	GTGTGATGGGGGACACCTGTAATCCAGCTATGCTGAGGCTGAGGCACGAGATCACTT	946
Db	118165	ATGTGGTGGCAGCGCGCTGTATAGTCCAGCTACCCAGAGGCTGAGACGAGATCGCTT	1182244
QY	947	GAACCTTGAGGCGGAGGTTGCACTGAGGCGCCGAGATCAATCAGCGCTTCACTGGGC	1006
Db	118225	GAACCCAGAGGCGGAAAGTTGCACTGAGGCTGAGATCAAGCCCACTGCACTCAGCTGGGT	1182844
QY	1007	GACAGAGCAAGACTCTGTCTCAAAATTAATTAATTAACA	1045
Db	118285	GATGAGCAAGACTCTCATCTTAAAAAAGAAAAA	118323
RESULT 13			
ABK83562/C	ABK83562/C		
ID	ABK83562 standard; cDNA, 139904 BP.		
XX	ABK83562;		
XX	14-AUG-2002 (first entry)		
DE	Human cDNA differentially expressed in granulocytic cells #133.		
XX	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
XX	viral infection; parasitic infection; protozoal infection;		
XX	fungal infection; sterile inflammatory disease; psoriasis;		
XX	rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;		
XX	cardiac reperfusion injury; renal reperfusion injury; AIDS;		
XX	adult respiratory distress syndrome; inflammatory bowel disease;		
XX	Crohn's disease; ulcerative colitis; peridontal disease;		
XX	granulocyte activation; chronic inflammation; allergy.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	WO200228999-A2.		
PN	11-APR-2002.		
PD	03-OCT-2001; 2001WO-US030821.		
XX	03-OCT-2000; 2000US-0237189P.		
XX	(GENE-) GENE LOGIC INC.		
PA	Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;		
PI	WPI, 2002-435328/46.		
XX	Detecting granulocyte activation by detecting differential expression of		
XX	genes associated with granulocyte activation, which serves as diagnostic		
XX	markers that is useful for monitoring disease states and drug toxicity.		
XX	Claim 1, SEQ ID NO 133; 114pp; English.		
XX	The invention relates to detecting (M1) granulocyte (GC) activation		
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by		
CC	DNA chip analysis as given in the specification, and comparing the		
CC	expression level to an expression level in an unactivated GC, where		
CC	differential expression of Gs is indicative of GCA. Also included are		
CC	modulating (M2) Gs by contacting GC with an agent that alters the		
CC	expression of at least one gene in Gs; (2) screening (M3) for an agent		
CC	capable of modulating GCA or an inflammation (especially chronic) in a		
CC	tissue, an allergic response in a subject, exposure of a subject to a		
CC	pathogen or sterile inflammatory disease using the gene expression		
CC	profile; (3) detecting (M4) an inflammation (especially chronic) in a		
CC	tissue, an allergic response (M4) an inflammation (especially chronic) in a		
CC	pathogen or sterile inflammatory disease, by detecting the level of		
CC	expression in a sample of the tissue of gene(s) from Gs, where the level		
CC	of expression of the gene is indicative of inflammation; (4) treating		

[illegible]

QY	649	CGCCGGGCATGATCATCTTCTTGCATAGTATGGAACAATCCTAAAGCATCACA	708
Db	15449	CACCCGGC-----GATGTTTGCACACTTTAAACAGTAAACAAATCCCAAGTTTAA--	15501
QY	709	CTCAACCTTAAATATGCAAGTTTGGCCAGAGCACCTGGCTCATGCTGTATATTCACG	768
Db	15502	-----TAAATATATTAAGCTGGGGGGGGGGTGGCTCACACTGTATCCACG	15552
QY	769	ACTTTGGAGGCAAGGTGGGTGAATCATCTTGAAGCCAGAGATTTGAACCAAGCTGGCC	828
Db	15553	AGTTTGGAGGGCCAGAGCAGGTGATCATCTTGAGCTCGGTGTTCAAGACCAAGCTTGGCC	15612
QY	829	AACATGTGAAACCTGTCTTACTTAATAAAAAAAAAAAAAAAAAAAATCAGCCGGGT	888
Db	15613	AACATGTGAAACCTGTCTTACT-----AAAAAAAAAAATTAAGCCAGGT	15660
QY	889	GTGTGGGGGCACACTGTATCCAGCTATGCTGAGGCTGAAGCAGAGATCACTTGA	948
Db	15661	GTGTGGGCACACCTGTATCCAGCTATCTTGGAGGCTGAAGCAGAGAAATCACTTGA	15720
QY	949	AACCTGGAAGGCGAGGTTGCAGTGGGCGGAGATCATCAACGGCCTCCAGCTGGGGCA	1008
Db	15721	AACCGGGAGGCGAGGTTGCAGTGAGCGGAGATTATGCCACTCCACTCTGTGGGTGA	15780
QY	1009	CAGAGCAAGACTGTCTCAAAATTAATTAATTAACAAGAACACAGT	1058
Db	15781	AAGAGCAAGACTGTGTCTCAAAACATTTAAATTAATTAATTAAGT	15830

```
Search completed: February 25, 2005, 20:14:49
Job time : 2552.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 18:05:22 ; Search time 8524.79 Seconds
(without alignments)
11373.754 Million cell updates/sec

```

Title:      US-09-936-271B-13_COPY_6000_8000
Perfect score: 2001
Sequence:   1 gacaaccaagagccccaag.....tcgaggggggaaacttta 2001

```

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters:

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```

1:  GenEmbl1: *
2:  gb_ba: *
3:  gb_hvg: *
4:  gb_in: *
5:  gb_cm: *
6:  gb_ov: *
7:  gb_pat: *
8:  gb_ph: *
9:  gb_pl: *
10: gb_pr: *
11: gb_ro: *
12: gb_sts: *
13: gb_ey: *
14: gb_un: *
15: gb_vt: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2001	100.0	11570	6	CQ788219	Sequence
2	2001	100.0	11570	6	CQ874885	Sequence
3	2001	100.0	11570	6	CQ874961	Sequence
4	2001	100.0	11570	9	AF135028	Homo sapi
5	2001	100.0	107487	9	AC011483	Homo sapi
6	2001	100.0	230000	9	AF243527	Homo sapi
7	1968.6	98.4	217346	2	AC027602	Homo sapi
8	1927.2	96.3	200792	2	AC130782	Pan trogl
9	286.4	14.3	142003	2	AC098680	Homo sapi
10	276.8	13.8	177787	9	AL133961	Homo sapi
11	274.4	13.7	13411	9	HSDJ71H1.9	Human DNM
12	264	13.2	161014	9	AC087591	Homo sapi
13	264	13.2	177876	2	AC025163	Homo sapi
14	262.4	13.1	92882	9	AL355815	Human DNM
15	261.8	12.9	85304	9	HS227L5	Human DNM
16	258.4	12.9	177562	9	HS269M15	Human DNM
17	257.8	12.9	70356	9	AL353794	Human DNM
18	257.8	12.9	163330	2	AC016011	Homo sapi
19	255.8	12.8	154664	9	AC025362	Homo sapi

ALIGNMENTS

C	20	255.8	12.8	160426	2	AC087657	Homo sapi
C	21	255.8	12.8	162354	2	AC073239	Homo sapi
C	22	255.8	12.8	179219	5	AC016489	Homo sapi
C	23	255.2	12.8	74705	5	AC116539	Homo sapi
C	24	255.2	12.8	161133	2	AC019295	Homo sapi
C	25	255.2	12.8	163690	2	AC083781	Homo sapi
C	26	255.2	12.8	195478	5	AC010259	Homo sapi
C	27	253.8	12.7	136649	9	HS016358	Homo sapi
C	28	253.4	12.7	160420	9	AC022023	Homo sapi
C	29	253.4	12.7	162609	9	AC021171	Homo sapi
C	30	253.4	12.7	179009	9	AL355543	Human DNA
C	31	253.2	12.7	173840	9	AC092445	Homo sapi
C	32	253.2	12.7	206819	2	AC016029	Homo sapi
C	33	253	12.6	193772	9	AC007073	Homo sapi
C	34	252.6	12.6	52316	6	BD012166	Isolation
C	35	252.6	12.6	52316	6	BD128650	Method fct
C	36	252.6	12.6	122302	9	AC003982	Homo sapi
C	37	252.6	12.6	163453	9	AC095851	Homo sapi
C	38	252.6	12.6	169544	9	AC008573	Homo sapi
C	39	252.6	12.6	179655	2	AC010246	Homo sapi
C	40	252.6	12.6	197810	1	AC090071	Homo sapi
C	41	252.6	12.6	220384	2	AC078926	Homo sapi
C	42	252.6	12.6	253038	2	AC008930	Homo sapi
C	43	251.4	12.6	33768	5	AC078783	Homo sapi
C	44	250.6	12.5	103585	9	AC113614	Homo sapi
C	45	250.6	12.5	180939	2	AL359620	Homo sapi

RESULT

LOCUS	CQ788219	11570 bp	DNA	linear	PAT 24-MAR-2004
DEFINITION	Sequence 2 from Patent WO2004021008.				
ACCESSION	CQ788219				
VERSION	CQ788219.1	GI:45723068			

REFERENCE

TITLE Methods for detecting breast and ovarian cancer
JOURNAL Patent: WO 2004021008-A 2 11-MAR-2004;
MOUNT SINAI HOSPITAL (CA); Yousef, George (CA)

ORIGIN

[illegible]

OY	241	TGAGGACACCTCTCTTTATTTAGCAGAGATACA	CACTGAGTGGCAACTGGTAACTATGAGGC	300
Db	6240	TGAGGACACCTCTCTTTATTTAGCAGAGATACA	CACTGAGTGGCAACTGGTAACTATGAGGC	6239
OY	301	GTTCGCAAAATCTGGAATCCAGCAATTCGCAAG	CAGTCAAGACCCCTGTTCTCAAGA	360
Db	6300	GTTCGCAAAATCTGGAATCCAGCAATTCGCAAG	CAGTCAAGACCCCTGTTCTCAAGA	6359
OY	361	GCTCATACCCCTAGAGTATGAGGTTTAACTAA	ATAAGTCGAGTGTCTTATATGTCATTTCC	420
Db	6360	GCTCATACCCCTAGAGTATGAGGTTTAACTAA	ATAAGTCGAGTGTCTTATATGTCATTTCC	6419
OY	421	CAGTTTTTTAGTAGCCACATTAAAA	CAGGTAAAAAGGCTGGGCGCAGTGGCTACA	480
Db	6420	CAGTTTTTTAGTAGCCACATTAAAA	CAGGTAAAAAGGCTGGGCGCAGTGGCTACA	6479
OY	481	GTAATCCAGCACTTTGGGAGGCTGAGGACG	AGATACCTTTTGTCAGAGATTGAGA	540
Db	6480	GTAATCCAGCACTTTGGGAGGCTGAGGACG	AGATACCTTTTGTCAGAGATTGAGA	6539
OY	541	CTAGCCTGGCCAAACATGGCGAAAC	CTCTGTCTTAAAAAAAATCAAAAATTAACCTGGC	600
Db	6540	CTAGCCTGGCCAAACATGGCGAAAC	CTCTGTCTTAAAAAAAATCAAAAATTAACCTGGC	6599
OY	601	ATGTGGCGGGCGCCTGTAATCTCAGCTGCT	CAGAGGGCCGAGACACAGAAATCACTTAA	660
Db	6600	ATGTGGCGGGCGCCTGTAATCTCAGCTGCT	CAGAGGGCCGAGACACAGAAATCACTTAA	6659
OY	661	ACCCAGAGAGTGGAGGTTGACGTAGC	AGATATGCTGCATCTCAACCTGGGAGAC	720
Db	6660	ACCCAGAGAGTGGAGGTTGACGTAGC	AGATATGCTGCATCTCAACCTGGGAGAC	6719
OY	721	AGAGTGAACACTTGTCTCAAAAAGAAAA	AAAAAAACAAGTAAAAAGAAACAAGTGAAG	780
Db	6720	AGAGTGAACACTTGTCTCAAAAAGAAAA	AAAAAAACAAGTAAAAAGAAACAAGTGAAG	6779
OY	781	TTAACTTAAATAAACCAATGTATCCCAAT	CAATCATTTCAAAAGTGAATTAATATAA	840
Db	6780	TTAACTTAAATAAACCAATGTATCCCAAT	CAATCATTTCAAAAGTGAATTAATATAA	6839
OY	841	ACAATTATGAATGAGATACTTTACATCTTT	CTTCTGTTCATATATTAAGCTTTTGAAGT	900
Db	6840	ACAATTATGAATGAGATACTTTACATCTTT	CTTCTGTTCATATATTAAGCTTTTGAAGT	6899
OY	901	GAGTATATATGTTATGCTGACAGCACATCT	CAATTTGAGCTAGCTACATTCAGGTGCTC	960
Db	6900	GAGTATATATGTTATGCTGACAGCACATCT	CAATTTGAGCTAGCTACATTCAGGTGCTC	6959
OY	961	AGTAGCCCAATGTGGCTGACAGTTATCTG	TATTTGATGTGCAGGATCTTAGAGGAAATC	1020
Db	6960	AGTAGCCCAATGTGGCTGACAGTTATCTG	TATTTGATGTGCAGGATCTTAGAGGAAATC	7019
OY	1021	AGGGCTGTTTTGTATGTGTTGGGACAG	TGTGTCACGTCACTAAAGATACCATATCT	1080
Db	7020	AGGGCTGTTTTGTATGTGTTGGGACAG	TGTGTCACGTCACTAAAGATACCATATCT	7079
OY	1081	GGCACTCCGTGTATCAGATGTCAAGTTT	TGCGAGTTTCAAGCGTGTGTAGTAAAGTGC	1140
Db	7080	GGCACTCCGTGTATCAGATGTCAAGTTT	TGCGAGTTTCAAGCGTGTGTAGTAAAGTGC	7139
OY	1141	TTGTTTCAACAAAATCTGTAAATATGAC	AGTTTCTTAGCAATGCTCGTAAATATCTTGA	1200
Db	7140	TTGTTTCAACAAAATCTGTAAATATGAC	AGTTTCTTAGCAATGCTCGTAAATATCTTGA	7199
OY	1201	GGAAAGAAAGAAATCTGGTAGGTATTTT	TACAGAAATATTTAATACAGGGAGTTA	1260
Db	7200	GGAAAGAAAGAAATCTGGTAGGTATTTT	TACAGAAATATTTAATACAGGGAGTTA	7259
OY	1261	ATTGCAAAAGCTGTGGAAGGGCTGAGAA	CAAAAGTTAAAAATTAATAATCTCTGTGCT	1320
Db	7260	ATTGCAAAAGCTGTGGAAGGGCTGAGAA	CAAAAGTTAAAAATTAATAATCTCTGTGCT	7319

QY	1321	TAGATCTGCCTAATATGGGCAATTTGACAGAGTGTAAAGTTAAACCCCAAAATTAAC	1380
Db	7320	AAGAAATCTGCATTAATAGGCGAATTTGAGAGTGTAAAGTTAAACCCCAAAATTAAC	7379
QY	1381	ATGCTTTTAGGATAGTAACATATAGGCGCAATATTCAAAAGGTGGTCAGGGGAGCCTC	1440
Db	7380	ATGCTTTTAGGATAGTAACATATAGGCGCAATATTCAAAAGGTGGTCAGGGGAGCCTC	7439
QY	1441	CTTGAGAGGCTGGCAATTTGACACAGATGATACAACAAGAAGCTTAACTCGTGAAGT	1500
Db	7440	CTTGAGAGGCTGGCAATTTGACACAGATGATACAACAAGAGCTTAACTCGTGAAGT	7499
QY	1501	TTAAGGGGAAAGAAAGGACGTCGCAAAAGCCTGAGGCAGTAAAGAAATTTGGCTGATTC	1560
Db	7500	TTAAGGGGAAAGAAAGGACGTCGCAAAAGCCTGAGGCAGTAAAGAAATTTGGCTGATTC	7559
QY	1561	AAAGAAAGAGGAAACCAATATGCAATCTGAGAACAAATATGGGGGCAACGTAGAAAGTG	1620
Db	7560	AAAGAAAGAGGAAACCAATATGCAATCTGAGAACAAATATGGGGGCAACGTAGAAAGTG	7619
QY	1621	ACGCTGAGAGGTATGAGCAGGGGCGAATGCTCTCGAATATTTCTTGCTCAACAACAGA	1680
Db	7620	ACGCTGAGAGGTATGAGCAGGGGCGAATGCTCTCGAATATTTCTTGCTCAACAACAGA	7679
QY	1681	GCTTCCTATATGTTTAATGAAAGCTGATCTGTTGAGAAAGACAGAAATTTAAATCAAC	1740
Db	7680	GCTTCCTATATGTTTAATGAAAGCTGATCTGTTGAGAAAGACAGAAATTTAAATCAAC	7739
QY	1741	TGTTACATCAACGACGACCCCTCTCTTATTCAGGCTCCCAAGGAGCTTAGAAGGACGTA	1800
Db	7740	TGTTACATCAACGACGACCCCTCTCTTATTCAGGCTCCCAAGGAGCTTAGAAGGACGTA	7799
QY	1801	AGTTAAACAGCTCTCATTTAGCAGGGGTGTGTTTCAACAGTAGTTAGGAAGCTGGGATTT	1860
Db	7800	AGTTAAACAGCTCTCATTTAGCAGGGGTGTGTTTCAACAGTAGTTAGGAAGCTGGGATTT	7859
QY	1861	CAGGAGTACTCAGTCCCATAGGCTAATAAAAGCTCCCCCAAAATTTGACAAACCTGACAA	1920
Db	7860	CAGGAGTACTCAGTCCCATAGGCTAATAAAAGCTCCCCCAAAATTTGACAAACCTGACAA	7919
QY	1921	ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGCTTGAGCTTGAGGGGGGTGGG	1980
Db	7920	ATGCAACACCTCCCAAGCTCTCCCAATTTCTTCTGTGCTTGAGCTTGAGGGGGGTGGG	7979
QY	1981	TTGCGAGGGGGAAAACTTTTA 2001	
Db	7980	TTGCGAGGGGGAAAACTTTTA 8000	

REFERENCE	1
AUTHORS	Diamandis, E. P. and Petraki, C.
TITLE	Assay for detection of renal cell carcinoma
JOURNAL	Patent: WO 2004077060-A 3 10-SEP-2004; Mount Sinai Hospital (CA)
FEATURES	location/Qualifiers
SOURCE	1..11570 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"
ORIGIN	
Query Match	100.0%; Score 2001; DB 6; Length 11570;

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Diamondis, E. P.
TITLE Multiple marker assay for detection of ovarian cancer
JOURNAL Patent: WO 2004075713-A 5 10-SEP-2004;
MOUNT SINAI HOSPITAL CORPORATION (CA)
FEATURES
source 1. .11570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 2001; DB 6; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACACCAAGAGGCCCAAGAGTGTCCAGGTTCTTCTTGAATCCGACCTCTGTC 60
DB 6000 GACACCAAGAGGCCCAAGAGTGTCCAGGTTCTTCTTGAATCCGACCTCTGTC 6059
QY 61 CGCCTTCACCTTCTTCTCACTTCTCATTTGTTCCTGTTGACAGTGCATTCCTTAAG 120
DB 6060 CGCCTTCACCTTCTTCTCACTTCTCATTTGTTCCTGTTGACAGTGCATTCCTTAAG 6119
QY 121 TCTCCAGTCTGTAATATCAGCGTGTAAATGCAAAAAGTCCAGAGATCTTACCCGA 180
DB 6120 TCTCCAGTCTGTAATATCAGCGTGTAAATGCAAAAAGTCCAGAGATCTTACCCGA 6179
QY 181 GACAGATAGATGACACATGTTCTGCGCGGTGACAAAGCAGTGAAGACTCCTGTCAG 240
DB 6180 GACAGATAGATGACACATGTTCTGCGCGGTGACAAAGCAGTGAAGACTCCTGTCAG 6239
QY 241 TGAGGACACCTCTCTTATTCAGCAGATACACAGTGTGCACTCGGTAAATGAGAG 300
DB 6240 TGAGGACACCTCTCTTATTCAGCAGATACACAGTGTGCACTCGGTAAATGAGAG 6299
QY 301 GTTGCCAAATTCGAGATTCAGCAATTTGCCAAGACGTCAGACCCCTGTTCTCAGAG 360
DB 6300 GTTGCCAAATTCGAGATTCAGCAATTTGCCAAGACGTCAGACCCCTGTTCTCAGAG 6359
QY 361 GCTCATACCTTAGATAGTGTGTTAGTGAATATATGCTGAGCTTATGTCATTTTC 420
DB 6360 GCTCATACCTTAGATAGTGTGTTAGTGAATATATGCTGAGCTTATGTCATTTTC 6419
QY 421 CAGTTTTTATGATGACCATTTAAACAGGTAAAAAGGCTGGCGAGTGGCTCACACCT 480
DB 6420 CAGTTTTTATGATGACCATTTAAACAGGTAAAAAGGCTGGCGAGTGGCTCACACCT 6479
QY 481 GTTAATCCGACACTTTGGAGGCTGAGGCGAGGACGATCACCTTTTGTCAAGAGTTGAGA 540
DB 6480 GTTAATCCGACACTTTGGAGGCTGAGGCGAGGACGATCACCTTTTGTCAAGAGTTGAGA 6539
QY 541 CTAGCCTGGCCAAATGAGCAATCTGCTCTTAAAAAAATATCAAAAATTAGCCCTGAG 600
DB 6540 CTAGCCTGGCCAAATGAGCAATCTGCTCTTAAAAAAATATCAAAAATTAGCCCTGAG 6599
QY 601 ATGGTGGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGCGAGACCAAGAAATCACTTAA 660
DB 6600 ATGGTGGCGGCGCTGTATATCTCAGCTGCTCAGAGGCGCGAGACCAAGAAATCACTTAA 6659
QY 661 ACCCAAGAGGTGAGAGTTGAGAGTGTGATGATGTCCTCACTCCAACTGGAGAC 720
DB 6660 ACCCAAGAGGTGAGAGTTGAGAGTGTGATGATGTCCTCACTCCAACTGGAGAC 6719
QY 721 AGAGTGAACCTTTGTCTCAAAAAGAAAAAAACAAGTAAAAAGAAACAGGTGAAG 780
DB 6720 AGAGTGAACCTTTGTCTCAAAAAGAAAAAAACAAGTAAAAAGAAACAGGTGAAG 6779

QY 781 TTAACCTTAATAACCAATGATATCCCAATATCAATCATTTCAAGTATTAATATAA 840
DB 6780 TTAACCTTAATAACCAATGATATCCCAATATCAATCATTTCAAGTATTAATATAA 6839
QY 841 ACAATTAATGAATGAGATTAATTTACATTTCTTTCTGTTTCAATTAAGCTTTGAAAT 900
DB 6840 ACAATTAATGAATGAGATTAATTTACATTTCTTTCTGTTTCAATTAAGCTTTGAAAT 6899
QY 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTGATCAATTTGAGGCTC 960
DB 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACCTGATCAATTTGAGGCTC 6959
QY 961 AGTAGCAGATGTGGCTAGCACTTACTGATTTGATGACCGAGATCTAGAGGAAAGATC 1020
DB 6960 AGTAGCAGATGTGGCTAGCACTTACTGATTTGATGACCGAGATCTAGAGGAAAGATC 7019
QY 1021 AGGCTGTTTTGTATGCTTGGGAGGTTGTGACATGATTAAGATATCTAATATAG 1080
DB 7020 AGGCTGTTTTGTATGCTTGGGAGGTTGTGACATGATTAAGATATCTAATATAG 7079
QY 1081 GGCACCTCGGTGTAAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGATTAAGTGC 1140
DB 7080 GGCACCTCGGTGTAAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGATTAAGTGC 7139
QY 1141 TTGTTTCAACAAATCTGTATATGACATTTTCTAGCAAGTGTGTTAAATATCTTGA 1200
DB 7140 TTGTTTCAACAAATCTGTATATGACATTTTCTAGCAAGTGTGTTAAATATCTTGA 7199
QY 1201 GGAAGAAAAAGAAATCTGTATGTTATTTTACAGAGATATTTAATACAGGGATTA 1260
DB 7200 GGAAGAAAAAGAAATCTGTATGTTATTTTACAGAGATATTTAATACAGGGATTA 7259
QY 1261 ATTGCAAGCTGCTGGAAGGCTGAGAGAACAAAGTTAAAAATAAAACTCTGAGTGC 1320
DB 7260 ATTGCAAGCTGCTGGAAGGCTGAGAGAACAAAGTTAAAAATAAAACTCTGAGTGC 7319
QY 1321 AAGAAATCTGCATTAATAGGCAATTTACAGAGTGTATTAAGTAAATTAATTAAC 1380
DB 7320 AAGAAATCTGCATTAATAGGCAATTTACAGAGTGTATTAAGTAAATTAATTAAC 7379
QY 1381 ATGGTTTTAGATATGTAACATTAAGGCTTATTTCAAAAAGTGTGTCAAGGAGCTTC 1440
DB 7380 ATGGTTTTAGATATGTAACATTAAGGCTTATTTCAAAAAGTGTGTCAAGGAGCTTC 7439
QY 1441 CTTGAGAGGTGGCAATTTGAGCAGAGAAATGATGACACAAAGAGTAAATCTGTAAGT 1500
DB 7440 CTTGAGAGGTGGCAATTTGAGCAGAGAAATGATGACACAAAGAGTAAATCTGTAAGT 7499
QY 1501 TTAAGGGGAAAAAGGACGTCGAAAGGCCCTGAGGCAATTAAGAAATTTGGCTGATTC 1560
DB 7500 TTAAGGGGAAAAAGGACGTCGAAAGGCCCTGAGGCAATTAAGAAATTTGGCTGATTC 7559
QY 1561 AAAAGAAAGAGAAACCAATGCACTGAGAGAACAAAAGTGGGGCAACGTAAGAAAGT 1620
DB 7560 AAAAGAAAGAGAAACCAATGCACTGAGAGAACAAAAGTGGGGCAACGTAAGAAAGT 7619
QY 1621 ACGCTGAGAGTGTAGGCAAGGCGCAATGCTCTGCAAGTATTTCTTGTGTCAACAAACAGA 1680
DB 7620 ACGCTGAGAGTGTAGGCAAGGCGCAATGCTCTGCAAGTATTTCTTGTGTCAACAAACAGA 7679
QY 1681 GCTTCCCTATGTTCTTAATGAAAGCTGTATCTGTGTAAGAGACAGAAATTTAAATCAAC 1740
DB 7680 GCTTCCCTATGTTCTTAATGAAAGCTGTATCTGTGTAAGAGACAGAAATTTAAATCAAC 7739
QY 1741 TGTATCATTAACACAGACCTTTCTCTGTATTAAGGCTCCCAAGGATCTAAGAAAGACGTA 1800
DB 7740 TGTATCATTAACACAGACCTTTCTCTGTATTAAGGCTCCCAAGGATCTAAGAAAGACGTA 7799
QY 1801 AGTTAAACAAGCTCTCATTTAGCAGAGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAT 1860
DB 7800 AGTTAAACAAGCTCTCATTTAGCAGAGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAT 7859
QY 1861 CAGAGTACTCCAGTCCCATGTATGAAGCTCCGCCCAATTTGTACAAACCTGACAA 1920

Db 7660 CAGAGATCTCCATCCATGCTATGAAAAGCTCCCCCAATTGTCACAACTGACAA 7919

Qy 1921 ATGCAACACCTCCCAAGCTCTCCCATTTCTCTGTGCTGGGTGGGGGGGGG 1980

Db 7920 ATGCAACACCTCCCAAGCTCTCCCATTTCTCTGTGCTGGGTGGGGGGGGG 7979

Qy 1981 TTGGAGGGGGGAAACTTTTA 2001

Db 7980 TTGGAGGGGGGAAACTTTTA 8000

RESULT 4
AF135028 11570 bp DNA linear PRI 27-JUN-2000

LOCUS Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.
DEFINITION AF135028
ACCESSION AF135028
VERSION AF135028.1 GI:4589282
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 11570)
AUTHORS Yousef, G.M., and Diamandis, E.P.
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,
mapping, tissue expression, and hormonal regulation
JOURNAL J. Biol. Chem. 274 (53), 37511-37516 (1999)
MEDLINE 20076408
PUBMED 10608802

REFERENCE 2 (bases 1 to 11570)
AUTHORS Diamandis, E.P., Yousef, G.M., Luo, L.Y., Magklara, A., and Obiezu, C.V.
TITLE The new human kallikrein gene family: implications in
carcinogenesis
Trends Endocrinol. Metab. 11 (2), 54-60 (2000)
MEDLINE 21121728
PUBMED 10675891

REFERENCE 3 (bases 1 to 11570)
AUTHORS Yousef, G.M., Luo, L.Y., and Diamandis, E.P.
TITLE Direct Submision
Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada

FEATURES
SOURCE Location/Qualifiers
1. .11570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.3-q13.4"
join(2101. .2131,2210. .2293,4762. .5023,5763. .6019,
6105. .6238,11092. .11570)
/product="kallikrein-like protein 2 KLK-L2"
join(2221. .2293,4762. .5023,5763. .6019,6105. .6238,
11092. .11247)
/codon_start=1
/product="kallikrein-like protein 2 KLK-L2"
/protein_id="AA026429.1"
/db_xref="GI:4589283"
/translation="MATARPMMVNLALITALLGTEHYLANDVSCDPSNTVPS
GSNDLGAAGEDARSDSSRITNGSDCMHTOPKQALILRPNOIYCAAVLVHPWM
LITPAHCRKRYFRRLGHSISPYBSGQMFQGVKSIPIHFGSHPHSNDMLIKAN
RLRPITDVRPIVNSHCPKAGTCLVSGWGTTSPOVHPKVLQCLNISLVLSKRC
DAYRQIDITMFCADKAGRDSQSDSGSPVNCNGLQGLVSWGDYPCARPNRPGVYT
NLKFTKMIQSTIQANS"

ORIGIN
Query Match 100.0%; Score 2001; DB 9; Length 11570;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GACAACCAAGAGCCCAAGGTGAGTCCAGGTTCTTCTGATACGACCATCTGCG 60
|||||

Db 6000 GACAACCAAGAGCCCAAGGTGAGTCCAGGTTCTTCTGATACGACCATCTGCG 6059

Qy 61 CGCCTTCATCTTCTTCACTTCTCATTTGTTCTGTTGACAGTGCATCTCCCTAAG 120

Db 6060 CGCCTTCATCTTCTTCACTTCTCATTTGTTCTGTTGACAGTGCATCTCCCTAAG 6119

Qy 121 TCCCTCAGAGCTGAATATCAGCGGTGTAAGTACAGAAAAGTGCAGAGTGTACCCGA 180

Db 6120 TCCCTCAGAGCTGAATATCAGCGGTGTAAGTACAGAAAAGTGCAGAGTGTACCCGA 6179

Qy 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGCTCTGCCAG 240

Db 6180 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGCTCTGCCAG 6239

Qy 241 TGAGGACACCTCTCTTTATTACAGCATATACACTGAGTGCCTGCTGTTAATGAGC 300

Db 6240 TGAGGACACCTCTCTTTATTACAGCATATACACTGAGTGCCTGCTGTTAATGAGC 6299

Qy 301 GTTGCCTGAATTCGAGAAATCAGCAATTCGCAAGCAGTGAAGCCCTGTTCCACAGA 360

Db 6300 GTTGCCTGAATTCGAGAAATCAGCAATTCGCAAGCAGTGAAGCCCTGTTCCACAGA 6359

Qy 361 GCTCATACCTCTAGAGTATGATGTTTATGAGAAATATGCTGAGCTGCTTATGTCATTC 420

Db 6360 GCTCATACCTCTAGAGTATGATGTTTATGAGAAATATGCTGAGCTGCTTATGTCATTC 6419

Qy 421 CAGTTTTTATGATGCCACATTTAAACAGTTAAACAGTTAAACAGTTAAACAGTTAA 480

Db 6420 CAGTTTTTATGATGCCACATTTAAACAGTTAAACAGTTAAACAGTTAAACAGTTAA 6479

Qy 481 GTATATCCACACCTTTGGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAGAGTTTGA 540

Db 6480 GTATATCCACACCTTTGGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAGAGTTTGA 6539

Qy 541 CTAGCTTGCCCAACATGCGCAAACTGCTCTTAAACAAATATGACCTGCGC 600

Db 6540 CTAGCTTGCCCAACATGCGCAAACTGCTCTTAAACAAATATGACCTGCGC 6539

Qy 601 ATGCTGGCGGCGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACACAGATCATTTAA 660

Db 6600 ATGCTGGCGGCGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACACAGATCATTTAA 6659

Qy 661 ACCGAGAGGTGAGAGTTCAGTGAAGCTGAGATCGAGCACTGCTCAACTGGGAGAC 720

Db 6660 ACCGAGAGGTGAGAGTTCAGTGAAGCTGAGATCGAGCACTGCTCAACTGGGAGAC 6719

Qy 721 AGAGTGACATTTTGTCTCAAAAGAAAAAACAATGTAATTAATTAATTAATTAATTA 780

Db 6720 AGAGTGACATTTTGTCTCAAAAGAAAAAACAATGTAATTAATTAATTAATTAATTA 6779

Qy 781 TTAACCTTAATTAACCAATGTAATCCAAATACATCATTTCAAGTGTAAATTAATTA 840

Db 6780 TTAACCTTAATTAACCAATGTAATCCAAATACATCATTTCAAGTGTAAATTAATTA 6839

Qy 841 ACAATTAATGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 900

Db 6840 ACAATTAATGAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6899

Qy 901 GAGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

Db 6900 GAGTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6959

Qy 961 AGTAGCCACATGTTGGTGAAGTAACTGTAATTTGATGAGCAGATCTTAAGGAGAAATC 1020

Db 6960 AGTAGCCACATGTTGGTGAAGTAACTGTAATTTGATGAGCAGATCTTAAGGAGAAATC 7019

Qy 1021 AGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1080

Db 7020 AGGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7079

Qy 1081 GGCATCTCGTGTACAGATGTCAGTTTGGCAGTTTGGCAGGCTGTGTAGTTAAGTGT 1140

Db 7080 GGCATCTCGTGTACAGATGTCAGTTTGGCAGTTTGGCAGGCTGTGTAGTTAAGTGT 7139

Oy	1141	TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGGTAAATATCTTGA	1200
Db	7140	TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGGTAAATATCTTGA	7199
Oy	1201	GGAAGAAAAGAAATCTGGTAGTATTTTTCAGAGAAATATTTAATACAGGGATTA	1260
Db	7200	GGAAGAAAAGAAATCTGGTAGTATTTTTCAGAGAAATATTTAATACAGGGATTA	7259
Oy	1261	ATTGCAAAAGCTGTGGAAGGGCTGGAGAACAAAGTTAAAAATAAAAATCTGTGTCTC	1320
Db	7260	ATTGCAAAAGCTGTGGAAGGGCTGGAGAACAAAGTTAAAAATAAAAATCTGTGTCTC	7319
Oy	1321	AAGAATCTGCATTAATAGGCAATTTTCAGAGAGGTAAAGTTAACCCCAAAATTAAC	1380
Db	7320	AAGAATCTGCATTAATAGGCAATTTTCAGAGAGGTAAAGTTAACCCCAAAATTAAC	7379
Oy	1381	ATGTTTTAGATAGTAAACATTAAGGCCAATATTCAAAAAGTGTCTCAGGGAGCTC	1440
Db	7380	ATGTTTTAGATAGTAAACATTAAGGCCAATATTCAAAAAGTGTCTCAGGGAGCTC	7439
Oy	1441	CTTGAGAGGTGTGCATTTGAGCAGAGATGATGACACAAAGACTTAACTGTGAAGT	1500
Db	7440	CTTGAGAGGTGTGCATTTGAGCAGAGATGATGACACAAAGACTTAACTGTGAAGT	7499
Oy	1501	TTAAGGGGAAAAGAAAAGGCACGTGCAAAAGGCCCTGAGGCAGTTAAGAAATTTGGCTGATTC	1560
Db	7500	TTAAGGGGAAAAGAAAAGGCACGTGCAAAAGGCCCTGAGGCAGTTAAGAAATTTGGCTGATTC	7559
Oy	1561	AAAGAAGAAAGAAAACCAATGCACTTGAGAAACAAAAGTGGGGGCAACAGTAGAAAGTG	1620
Db	7560	AAAGAAGAAAGAAAACCAATGCACTTGAGAAACAAAAGTGGGGGCAACAGTAGAAAGTG	7619
Oy	1621	ACGCTGAGGTGTAGGCAAGGGGCCAATGCTCTGCAAGTATTTCTTGTCACCAACACAGA	1680
Db	7620	ACGCTGAGGTGTAGGCAAGGGGCCAATGCTCTGCAAGTATTTCTTGTCACCAACACAGA	7679
Oy	1681	GCTTCCCTATGTTCTAATGAAGCTGTATCTGTTGAGAGAACAGAAATTTAAATCAAAAC	1740
Db	7680	GCTTCCCTATGTTCTAATGAAGCTGTATCTGTTGAGAGAACAGAAATTTAAATCAAAAC	7739
Oy	1741	TGTTACATCAACACAGACCCCTTCTCTGTATTCAGGCTCCCAAGGATCTAGAGAGACGTA	1800
Db	7740	TGTTACATCAACACAGACCCCTTCTCTGTATTCAGGCTCCCAAGGATCTAGAGAGACGTA	7799
Oy	1801	AGTTAACAAAGCTCATTAAGCAGGGTGTGTATTCACAGTATGTAGAACTGTGGGACTT	1860
Db	7800	AGTTAACAAAGCTCATTAAGCAGGGTGTGTATTCACAGTATGTAGAACTGTGGGACTT	7859
Oy	1861	CAGAGTACTCCAGTCCCAAGGCTATGAAAAGCTCCCCCAAAATTTGTACAACTGACAA	1920
Db	7860	CAGAGTACTCCAGTCCCAAGGCTATGAAAAGCTCCCCCAAAATTTGTACAACTGACAA	7919
Oy	1921	ATGCACAACCTTCCCAAGCTTCTCTCTGTGTGCTGTGCTGGGTGTGGGGGGGTGGG	1980
Db	7920	ATGCACAACCTTCCCAAGCTTCTCTCTGTGTGCTGTGCTGGGTGTGGGGGGGTGGG	7979
Oy	1981	TTTGCAGAGGGGAAAACCTTTTA	2001
Db	7980	TTTGCAGAGGGGAAAACCTTTTA	8000
RESULT 5			
ACOL1483/c	ACOL1483	107487 bp	DNA linear PRI 01-JUL-2002
LOCUS	ACOL1483		
DEFINITION	Homio sapiens chromosome 19 clone CTB-147C22, complete sequence.		
ACCESSION	ACOL1483		
VERSION	ACOL1483.7	GI:21637461	
KEYWORDS	HTG.		
SOURCE	Homio sapiens (human)		
ORGANISM	Homio sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		

Query Match	100.0%;	Score 2001;	DB 9;	Length 107487;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2001;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
1	GACAAACCAAGAGCCCCCAAGAGTGAAGTGCAGGTTCTTCTTGAATACCAACCATCTGAC	60		
54220	GACAAACCAAGAGCCCCCAAGAGTGAAGTGCAGGTTCTTCTTGAATACCAACCATCTGAC	54161		
61	CGCCTTCATCTTCTTCACATCTCTCATCTGTTCTGTTGACATGACATTCCTTAAG	120		
54160	CGCCTTCATCTTCTTCACATCTCTCATCTGTTCTGTTGACATGACATTCCTTAAG	54101		
121	TCTCTCAGTCTGTAATATGAGGTGTCTAAGTGAAGAAAAGTGCAGAGATGCTTACCCGA	180		
54100	TCTCTCAGTCTGTAATATGAGGTGTCTAAGTGAAGAAAAGTGCAGAGATGCTTACCCGA	54041		
181	GACAGATAGATGACACCATGTTCTGGGCGGTGACAAACAGATGAGATCTCTGCACAG	240		
54040	GACAGATAGATGACACCATGTTCTGGGCGGTGACAAACAGATGAGATCTCTGCACAG	53981		
241	TGAGGACACCTCTCTTATTCACAGATACACACTGAGTGCACATCTCGTAAATGAGAC	300		
53980	TGAGGACACCTCTCTTATTCACAGATACACACTGAGTGCACATCTCGTAAATGAGAC	53921		
301	GTTGCAAAATTCGAGATTCAGCAATTCGCAAGACATGTCAGACCCCTGTTCTACAGA	360		
53920	GTTGCAAAATTCGAGATTCAGCAATTCGCAAGACATGTCAGACCCCTGTTCTACAGA	53861		
361	GCTCATACCTTAAGATAGTGTGTTTAAGTAAATTAATCTGAGTCTTAATGATCTTTC	420		
53860	GCTCATACCTTAAGATAGTGTGTTTAAGTAAATTAATCTGAGTCTTAATGATCTTTC	53801		
421	CAGTTTTTTAGTACCAATTAACAGGTAAACAAAAGGTGGGCGCAGTGGCTCACACT	480		
53800	CAGTTTTTTAGTACCAATTAACAGGTAAACAAAAGGTGGGCGCAGTGGCTCACACT	53741		
481	GTAATCCAGACACTTGGGAGGCTGAGGACGACGATCACTTTGGTCAAGATTTGAGA	540		

Db	53740	GTATATCCAGACCTTTGGAGGCTGAGGACGAGATCACTTTGGTCAGAGATTGAA	53681
Oy	541	CTACCTTGCCCAATGCGGAACTGTGTCTTAAAAAAAATACAAAATTAGCTGCG	600
Db	53680	CTACCTTGCCCAATGCGGAACTGTGTCTTAAAAAAAATACAAAATTAGCTGCG	53621
Oy	601	ATGTGGCCGGCCGCTTAAATTCAGTGTCTCAGAGAGCCGAGACCAAGATCACTTAA	660
Db	53620	ATGTGGCCGGCCGCTTAAATTCAGTGTCTCAGAGAGCCGAGACCAAGATCACTTAA	53561
Oy	661	ACCCAGAGAGTGGAGTTGCGAGTGAATCGTGCCTCACTCACTCAACTGGAGAC	720
Db	53560	ACCCAGAGAGTGGAGTTGCGAGTGAATCGTGCCTCACTCACTCAACTGGAGAC	53501
Oy	721	AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGTGAAG	780
Db	53500	AGAGTGACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAGTGAAG	53441
Oy	781	TTAACTTTAATTAACCCCAATGTATCCCAATACATCTTTCAAAGTGAATTATTA	840
Db	53440	TTAACTTTAATTAACCCCAATGTATCCCAATACATCTTTCAAAGTGAATTATTA	53381
Oy	841	ACAATTATGATGATGATTAATTCAATCTTTCTGTGTTTCAATTAAGCTTGAAGT	900
Db	53380	ACAATTATGATGATGATTAATTCAATCTTTCTGTGTTTCAATTAAGCTTGAAGT	53321
Oy	901	GAGTATATATGTTATGTCGACAGACATCTCAATTGSACTAGTCAATTCAGTGTCT	960
Db	53320	GAGTATATATGTTATGTCGACAGACATCTCAATTGSACTAGTCAATTCAGTGTCT	53261
Oy	961	AGTAGCCACATGTGGCTAGCACTTACTGTATGAGTGGACGGATCTAAGGAAAGATC	1020
Db	53260	AGTAGCCACATGTGGCTAGCACTTACTGTATGAGTGGACGGATCTAAGGAAAGATC	53201
Oy	1021	AGGCTCTTTTGTATGTTGGGAGGTTGGACCTGATPAABATACCATCTAATAGG	1080
Db	53200	AGGCTCTTTTGTATGTTGGGAGGTTGGACCTGATPAABATACCATCTAATAGG	53141
Oy	1081	GGCACTCCGTGTACAGATGTCAAGTTTGGCAGTTTCAAGCCGTGTGATTAAGTGT	1140
Db	53140	GGCACTCCGTGTGTACAGATGTCAAGTTTGGCAGTTTCAAGCCGTGTGATTAAGTGT	53081
Oy	1141	TTGTTTCAACAAATCTGTAAATGTACAGTTTCTAGCAAGTCTGTAAATATCTTGA	1200
Db	53080	TTGTTTCAACAAATCTGTAAATGTACAGTTTCTAGCAAGTCTGTAAATATCTTGA	53021
Oy	1201	GGAGGAAAAAGAAATCTGTGTAGTATTTTTCAGAGAAATTTTAATACAGGGGATTA	1260
Db	53020	GGAGGAAAAAGAAATCTGTGTAGTATTTTTCAGAGAAATTTTAATACAGGGGATTA	52961
Oy	1261	ATTGCAAAAGCTGTGTAAGGCTGAGAGCAAAAGTTAAAAAATAAAAAATCTGTGTCT	1320
Db	52960	ATTGCAAAAGCTGTGTAAGGCTGAGAGCAAAAGTTAAAAAATAAAAAATCTGTGTCT	52901
Oy	1321	AAGAAATCTGTCAATAATAGGCCAATTCAGAGTGTAAAGTTAAACCCCAAAATAAAC	1380
Db	52900	AAGAAATCTGTCAATAATAGGCCAATTCAGAGTGTAAAGTTAAACCCCAAAATAAAC	52841
Oy	1381	ATGTTTTTAAAGATAGTAAACAAATTAAGGCCAATATTCAAAAAGTGTCAAGGAGCTTC	1440
Db	52840	ATGTTTTTAAAGATAGTAAACAAATTAAGGCCAATATTCAAAAAGTGTCAAGGAGCTTC	52781
Oy	1441	CTTGGAAGGTGTGCAATTTAGCAGAGATAGTGAACAAGAAAGCTAAACTGTGAAGT	1500
Db	52780	CTTGGAAGGTGTGCAATTTAGCAGAGATAGTGAACAAGAAAGCTAAACTGTGAAGT	52721
Oy	1501	TTAAAGGGAAGAAAGAAAGCACTGTCAAAAGCCCTGTAGGCACTAAGGAATTTGGCTGATTC	1560
Db	52720	TTAAAGGGAAGAAAGAAAGCACTGTCAAAAGCCCTGTAGGCACTAAGGAATTTGGCTGATTC	52661
Oy	1561	AAAGAAAGAGGAAACCAATGCACTGTGAGAAACAAAGTGGGGGCAACGTAGAAAGTGT	1620

Db	52660	AAAGAAAGAGGAAACCAATGCACTGGAGAACAAAGTGGGGGCAACAGTGAAGG	52601
Qy	1621	ACGCTGAAGGTGTATGGCAGAGGGGGAATGCTCTTCGCAAGTATTTTGTGTACCCAAACACAA	1680
Db	52600	ACGCTGAAGGTGTATGGCAGAGGGGGAATGCTCTTCGCAAGTATTTTGTGTACCCAAACACAA	52541
Qy	1681	GCTTCCTTATGTTCTATATGGAAGCTGTATCTGTGTGAGGAAGACAGAAATTTAAATCAAA	1740
Db	52540	GCTTCCTTATGTTCTATATGGAAGCTGTATCTGTGTGAGGAAGACAGAAATTTAAATCAAA	52488
Qy	1741	TGTTCATCAACAAGCACCCTTCTCTGTATTCAGGCTCCCAAGGATCTTGAAGACGTA	1800
Db	52480	TGTTCATCAACAAGCACCCTTCTCTGTATTCAGGCTCCCAAGGATCTTGAAGACGTA	52422
Qy	1801	AGTTAACAGCTCTCATTAGCAGAGGTGTGTATTCAACAGTATGAAAGTGGGGATT	1860
Db	52420	AGTTAACAGCTCTCATTAGCAGAGGTGTGTATTCAACAGTATGAAAGTGGGGATT	52361
Qy	1861	CAGAGATACCTCCAGTCCCATAGGTATGAAAGCTCCCAATTTGTACAAACCTGACAA	1920
Db	52360	CAGAGATACCTCCAGTCCCATAGGTATGAAAGCTCCCAATTTGTACAAACCTGACAA	52301
Qy	1921	ATGCACACACTCCCAAGCTCTCCCAATTTCTTCTGTGTGCGCTGGTGTGGGGGGTGG	1980
Db	52300	ATGCACACACTCCCAAGCTCTCCCAATTTCTTCTGTGTGCGCTGGTGTGGGGGGTGG	52241
Qy	1981	TTGCGAGGGGGAACCTTTTA 2001	
Db	52240	TTGCGAGGGGGAACCTTTTA 52220	
RESULT 6			
AF243527/c			
LOCUS	AF243527	230000 bp	DNA linear PRI 21-NOV-2000
DEFINITION	Homo sapiens serine protease gene cluster, complete sequence.		
ACCESSION	AF243527		
VERSION	AF243527.1	GI:11244757	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 230000)		
AUTHORS	Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,		
	Moss, P., Papenf, B. and Wang, K.		
TITLE	Sequencing and expression analysis of the serine protease gene		
	cluster located in chromosome 19q13 region		
JOURNAL	Gene 257 (1), 119-130 (2000)		
MEDLINE	20510030		
PUBMED	11054574		
REFERENCE	2 (bases 1 to 230000)		
AUTHORS	Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,		
	Moss, P., Papenf, B. and Wang, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,		
	Bothell, WA 98021, USA		
FEATURES	Location/Qualifiers		
source	1..230000		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="19"		
	/map="19q13"		
	complement(<6963..>11517)		
gene	/gene="KLK1"		
mRNA	complement(join(<6963..7118,7668..7804,7923..8212,		
	9482..9641,11472..>11517))		
	/gene="KLK1"		
	/product="renal kallikrein"		
	complement(join(6963..7118,7668..7804,7923..8212,		
	9482..9641,11472..11517))		
	/gene="KLK1"		
	/note="serine protease"		
CDS			

/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCALSLGCGGAPPTISRIYGVGECRQHSQPMQVALYH
FTFQCGGLVHPOWLTAAHCTSDNYQMLGHNHFDENTAQFVHVESFPHFEN
MSLEHNTQADEYSDHDLMLRLTEPADITIDAVVGLPTEPEVSGTCLASGWS
IEBENSFPDLOQVDLKLINDECKAHQKVTDFMLCVGLGEGKDTCVGDSGGPL
MCGVLQGVTSWYPCGTPNKPSVAVRVLSYKWIETIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15563,19204..>192461))
/product="ACO protease"
/complement(join(13552..13704,14377..14530,14678..14917,
15416..15563,19204..192461))
/note="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MMFLTLSPFLASTAAQODKLLGDEGCAPHSQPMQVALYHGR
FNCGLSLSPHWLSAAHCSQRMVRVLGSHNLRKRDGPOLRTTSRVI DHPYEARS
HNDIMLRLVNPBARLNPOVRPAVLPTPCDHGEACVSGKGIPLSSSPDITLHC
ANISITSDTSCDSKSYPRGLNTVTCAGABRGASCEGDSGGPLVCGILQGVTSWGD
VPCDNTTKPGVYTKCVHLEWIRETKRN"
/codon_start=1
/product="serine protease"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MMVVFVFLTSVTWIGAAPLILSRIVGECRQHSQPMQVALYH
RGAVCGVTVHPOWLTAAHCTIRKNSVILLGHSILFHPEDTGOVROVSHSPHLYD
MELLNKRLRPGDSSHDMLRLSEPAELTDAVKYMDLPTQEPALGTTCTYASGMSI
EPEFLTPKQLQVDLHVISNDYCAQVHPQVTKFMLCAGRWTGSKSTCSGDSGGPLV
CNGVLQGVTSWSEPCALPERPSLYTKVHYRKMIDTIIVANP"
/codon_start=1
/product="serine protease"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MMDIVLSTALSVGCGAVPLIOSRIVGECRQHSQPMQVALYH
HGMHAGGVTVHPOWLTAAHCTKRSQVWLGHNHFPEDTGOVROVSHSPHLYN
MSLTKQSLRPEDSSHDMLRLSEPAELTDAVKYMDLPTQEPALGTTCTYASGMSI
EPEFLTPKQLQVDLHVISNDYCAQVHPQVTKFMLCAGRWTGSKSTCSGDSGGPLV
CNGVLQGVTSWSEPCALPERPAVTVKHYRKMIDTIIVANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..>984051))
/product="serine protease"
/complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..984051))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPMGWFLGYLLGVAGSLVSGSCQIINGEDCSPHQPW
QVALVMENELFCGVLVHPOWLTAAHCFQNSYTIQIGLHSLPADQEPQSQWVEASTLS

/codon_start=1
/product="renal kallikrein"
/protein_id="AAG33353.1"
/db_xref="GI:11244758"
/translation="MMFLVLCALSLGCGGAPPTISRIYGVGECRQHSQPMQVALYH
FTFQCGGLVHPOWLTAAHCTSDNYQMLGHNHFDENTAQFVHVESFPHFEN
MSLEHNTQADEYSDHDLMLRLTEPADITIDAVVGLPTEPEVSGTCLASGWS
IEBENSFPDLOQVDLKLINDECKAHQKVTDFMLCVGLGEGKDTCVGDSGGPL
MCGVLQGVTSWYPCGTPNKPSVAVRVLSYKWIETIAENS"
complement(join(<13552..13704,14377..14530,14678..14917,
15416..15563,19204..>192461))
/product="ACO protease"
/complement(join(13552..13704,14377..14530,14678..14917,
15416..15563,19204..192461))
/note="serine protease"
/codon_start=1
/product="ACO protease"
/protein_id="AAG33354.1"
/db_xref="GI:11244759"
/translation="MMFLTLSPFLASTAAQODKLLGDEGCAPHSQPMQVALYHGR
FNCGLSLSPHWLSAAHCSQRMVRVLGSHNLRKRDGPOLRTTSRVI DHPYEARS
HNDIMLRLVNPBARLNPOVRPAVLPTPCDHGEACVSGKGIPLSSSPDITLHC
ANISITSDTSCDSKSYPRGLNTVTCAGABRGASCEGDSGGPLVCGILQGVTSWGD
VPCDNTTKPGVYTKCVHLEWIRETKRN"
/codon_start=1
/product="serine protease"
/protein_id="AAG33355.1"
/db_xref="GI:11244760"
/translation="MMVVFVFLTSVTWIGAAPLILSRIVGECRQHSQPMQVALYH
RGAVCGVTVHPOWLTAAHCTIRKNSVILLGHSILFHPEDTGOVROVSHSPHLYD
MELLNKRLRPGDSSHDMLRLSEPAELTDAVKYMDLPTQEPALGTTCTYASGMSI
EPEFLTPKQLQVDLHVISNDYCAQVHPQVTKFMLCAGRWTGSKSTCSGDSGGPLV
CNGVLQGVTSWSEPCALPERPSLYTKVHYRKMIDTIIVANP"
/codon_start=1
/product="serine protease"
/protein_id="AAG33356.1"
/db_xref="GI:11244761"
/translation="MMDIVLSTALSVGCGAVPLIOSRIVGECRQHSQPMQVALYH
HGMHAGGVTVHPOWLTAAHCTKRSQVWLGHNHFPEDTGOVROVSHSPHLYN
MSLTKQSLRPEDSSHDMLRLSEPAELTDAVKYMDLPTQEPALGTTCTYASGMSI
EPEFLTPKQLQVDLHVISNDYCAQVHPQVTKFMLCAGRWTGSKSTCSGDSGGPLV
CNGVLQGVTSWSEPCALPERPAVTVKHYRKMIDTIIVANP"
complement(join(<94602..94754,96027..96163,96247..96497,
96919..97081,98345..>984051))
/product="serine protease"
/complement(join(94602..94754,96027..96163,96247..96497,
96919..97081,98345..984051))
/codon_start=1
/product="serine protease"
/protein_id="AAG33357.1"
/db_xref="GI:11244762"
/translation="MATAGNPMGWFLGYLLGVAGSLVSGSCQIINGEDCSPHQPW
QVALVMENELFCGVLVHPOWLTAAHCFQNSYTIQIGLHSLPADQEPQSQWVEASTLS

194324..194589,197048..197204,197370..>197412)		/product="kallikrein-like 3"	
complement(join(190980..191129,191573..191709,194324..194589,197048..197204,197370..197412))		/note="serine protease"	
/codon_start=1		/product="kallikrein-like 3"	
/protein_id="PAG33362.1"		/db_xref="gi:11244767"	
/translation="MKRLGCLALLSLAGHMDTRAIAGEECPNSQPMQGLFHLT RLFGATLISDRMLTAHCKRKYLPVLRLEHMLKMEPEQLFRITDPEHPEFND LSANDHDDIMLRPLQARLSPAPVPLNSQTSVGMASVSPKLLFVY TLQCANISILENKICHNAYPGHISDSMLCAGLMEGSGSCQDSSGPLVNCGLTAGV SGAEPCSRPRRAVYTSVCHYLDIMQEIEMEN"			
Query Match	100.0%; Score 2001; DB 9; Length 230000;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	GACAAACGAGAGCCCAAGGTGAGTCCAGGTTCTTGATACCAACCATCTCTGC	60
DB	136548	GAGAACCAAGAGCCCCCAAGGTGAGTCCAGGTTCTTGATACCAACCATCTCTGC	136489
QY	61	CGCCTTCATCTTCTCCACTTCTCATTTGTTCCCTGTTGACAGTGACCTTCCCTAAG	120
DB	136488	CGCTTCATCTTCTCCACTTCTCATTTGTTCCCTGTTGACAGTGACCTTCCCTAAG	136429
QY	121	TCCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAAGGTGCGAGATCTTACCGA	180
DB	136428	TCCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAAGGTGCGAGATCTTACCGA	136369
QY	181	GACAGATGATGACCAATGTTTCCGCGGTGCAAAAGCGATGAGATCTCTGCGCAG	240
DB	136368	GACAGATGATGACCAATGTTTCCGCGGTGCAAAAGCGATGAGATCTCTGCGCAG	136309
QY	241	TGAGGACACCTCTCTTTATTCAGAGATACACATGAGTCCAACTCGGTAACTGAGC	300
DB	136308	TGAGGACACCTCTCTTTATTCAGAGATACACATGAGTCCAACTCGGTAACTGAGC	136249
QY	301	GTTGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCCGCTGTTCTACAGA	360
DB	136248	GTTGCCAAATTCAGAAATCCAGCAATTCGCAAGACAGTCCGCTGTTCTACAGA	136189
QY	361	GCTCATACCTAGAGTATGTTTATGATAAATATGCTGAGCTGTTATGTCATTTTC	420
DB	136188	GCTCATACCTAGAGTATGTTTATGATAAATATGCTGAGCTGTTATGTCATTTTC	136129
QY	421	CAGTTTCTAGTACCAATTAACAGGTAAACAGGTGAGGAGTGGCTGACACCT	480
DB	136128	CAGTTTCTAGTACCAATTAACAGGTAAACAGGTGAGGAGTGGCTGACACCT	136069
QY	481	GTAATCCAGCACTTTGGAGGCTGAGGAGGAGATCACTTTTGTCAGAGATTGAGA	540
DB	136068	GTAATCCAGCACTTTGGAGGCTGAGGAGGAGATCACTTTTGTCAGAGATTGAGA	136009
QY	541	CTAGCCTGGCCAAATGCGGAACTCTGTCTTAAACAAATTAACCTGGC	600
DB	136008	CTAGCCTGGCCAAATGCGGAACTCTGTCTTAAACAAATTAACCTGGC	135949
QY	601	ATGTTGCGGGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAGAACTCACTAA	660
DB	135948	ATGTTGCGGGGCGCTGTATCTCAGCTGCTCAGAGGCGGAGACACAGAACTCACTAA	135889
QY	661	ACCGAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCCAACTGGAGAC	720
DB	135888	ACCGAGAGGTGAGGTTGAGTGAAGTGAATGCTGCACTCACTCCAACTGGAGAC	135829
QY	721	AGAGTGAACCTTTGTCTCAAAAGAAAAAACAAGTAAAGAAACAGGTGAAG	780
DB	135828	AGAGTGAACCTTTGTCTCAAAAGAAAAAACAAGTAAAGAAACAGGTGAAG	135769
QY	781	TTAATCTTAATTAACCAATGTATCCAAATCAATCAATTTCAAGTGAATTAATATAA	840
DB	135768	TTAATCTTAATTAACCAATGTATCCAAATCAATCAATTTCAAGTGAATTAATATAA	135709
QY	841	ACAATTTGATGATGATACCTTTTACATTTCTTTCTTTTCAATTAAGCTTTGAAGT	900
DB	135708	ACAATTTGATGATGATACCTTTTACATTTCTTTCTTTTCAATTAAGCTTTGAAGT	135649
QY	901	GAGTATATATGTTATGCTGACAGACATCTCAATTTTGGATACATTTACAGTCTC	960
DB	135648	GAGTATATATGTTATGCTGACAGACATCTCAATTTTGGATACATTTACAGTCTC	135589
QY	961	AGTAGCCACATGTGGCTAGACAGTTACTGATTTGATGATGACAGATCTAGAGGAAATC	1020
DB	135588	AGTAGCCACATGTGGCTAGACAGTTACTGATTTGATGATGACAGATCTAGAGGAAATC	135529
QY	1021	AGGCTGTTTTTGTATGTTGGGCAAGTTTGTGCACTGATTAAGATACATATCTAATAG	1080
DB	135528	AGGCTGTTTTTGTATGTTGGGCAAGTTTGTGCACTGATTAAGATACATATCTAATAG	135469
QY	1081	GGCAGCTCGTTATACAGATGTCAGTTTGGCAGTTTCAAGCGGTGTAGTTAAGTGC	1140
DB	135468	GGCAGCTCGTTATACAGATGTCAGTTTGGCAGTTTCAAGCGGTGTAGTTAAGTGC	135409
QY	1141	TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGTGTAAATATCTTGA	1200
DB	135408	TTGTTTCAACAAATCTGTATATGACAGTTTCTAGCAAGTGTGTAAATATCTTGA	135349
QY	1201	GGAAGAAAAGAAATCTGTATGTTTCTTTTCAAGAGATTTTATACAGGGGATTA	1260
DB	135348	GGAAGAAAAGAAATCTGTATGTTTCTTTTCAAGAGATTTTATACAGGGGATTA	135289
QY	1261	ATTGCAAGGCTGTGGAAGGGCTGAGGAAACAAAGTTAAATAAATCTGCTGTC	1320
DB	135288	ATTGCAAGGCTGTGGAAGGGCTGAGGAAACAAAGTTAAATAAATCTGCTGTC	135229
QY	1321	AAGATCTGCAATTAATAGGCAATTTCAAGAGTGTGTAAGTTAACCCCAATAATAAC	1380
DB	135228	AAGATCTGCAATTAATAGGCAATTTCAAGAGTGTGTAAGTTAACCCCAATAATAAC	135169
QY	1381	ATGTTTATGATATGTAACATTAAGGCAATTTCAAAAGGTGTCAAGGGAGCTTC	1440
DB	135168	ATGTTTATGATATGTAACATTAAGGCAATTTCAAAAGGTGTCAAGGGAGCTTC	135109
QY	1441	CTTGGAGAGGTGCAATTTGAGCAGAGATGAGACCAAGAGCTAACTGTGTAAGT	1500
DB	135108	CTTGGAGAGGTGCAATTTGAGCAGAGATGAGACCAAGAGCTAACTGTGTAAGT	135049
QY	1501	TTAAGGGGAAAGAAAGGCACTGCAAAAGGCTGAGAGCGTAAAGAAATTTGGCTGATTC	1560
DB	135048	TTAAGGGGAAAGAAAGGCACTGCAAAAGGCTGAGAGCGTAAAGAAATTTGGCTGATTC	134989
QY	1561	AAAGAAAGAGAGAAACCAATGCACTGAGAACAAAGTGGGGCAACGTAGAAAGTG	1620
DB	134988	AAAGAAAGAGAGAAACCAATGCACTGAGAACAAAGTGGGGCAACGTAGAAAGTG	134929
QY	1621	ACCGTGAAGGTGAGGCAAGGGGGAATGCTCTGCAAGTATTTCTTGCTCAACAAACAGA	1680
DB	134928	ACCGTGAAGGTGAGGCAAGGGGGAATGCTCTGCAAGTATTTCTTGCTCAACAAACAGA	134869
QY	1681	GCTTCCCTATGTTCTAATGAACTGTATCTGTTGAGGAAAGACAAATTTAAATCAAC	1740
DB	134868	GCTTCCCTATGTTCTAATGAACTGTATCTGTTGAGGAAAGACAAATTTAAATCAAC	134809
QY	1741	TGTTACATCAACAGCAACCTTCTGTATTCAGGCTCCCAAGGATCTAGAGACGTA	1800
DB	134808	TGTTACATCAACAGCAACCTTCTGTATTCAGGCTCCCAAGGATCTAGAGACGTA	134749
QY	1801	AGTTAACAGCTCTCAATTAAGGAGGTGTGTTCAAGTATGTTGAGGAAAGGCGGGATTT	1860
DB	134748	AGTTAACAGCTCTCAATTAAGGAGGTGTGTTCAAGTATGTTGAGGAAAGGCGGGATTT	134689
QY	1861	CAGAGATCTCAGTCCCAATGAGCTATGAAAGGCTCCCAATTTGTAACAACTGACAA	1920
DB	134688	CAGAGATCTCAGTCCCAATGAGCTATGAAAGGCTCCCAATTTGTAACAACTGACAA	134629

QY 1921 ATGCAACACCTCCCGAGCTCTCCCATTTCTCTGTGCGCTGGGNGTGGGGGGG 1980
 |||||
 Db 134628 ATGCAACACTCCCGAGCTCTCCCATTTCTCTGTGCGCTGGGNGTGGGGGGG 134569
 |||||
 QY 1981 TTGCGAGGGGGAACCTTTTA 2001
 |||||
 Db 134568 TTGCGAGGGGGAACCTTTTA 134548
 |||||

RESULT 7
 AC027602/c 217346 bp DNA linear HTG 24-AUG-2002
 LOCUS Homo sapiens chromosome 19 clone RP11-795B6 map 19, WORKING DRAFT
 DEFINITION
 SQUENCE. 12 unordered pieces.
 AC027602
 AC027602.4 GI:11178143
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 Homo sapiens (human)
 SOURCE
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 217346)
 Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barna,N., Bascien,V., Beda,F.,
 Bogunlavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
 Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domingo,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
 Howland,J.C., Illey,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczký,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
 McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R.,
 Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 217346)
 Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Barna,N., Bascien,V., Beda,F., Bogunlavsky,L.,
 Boukhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
 Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
 FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
 Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L.,
 Illey,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Larocque,K.,
 Lamazares,R., Landers,T., Lehoczký,J., Levine,R., Lieu,C., Liu,G.,
 MacDonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K.,
 McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
 Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testave,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.
 Direct Submission

JOURNAL

COMMENT

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Nov 15, 2000 this sequence version replaced gi:11136831.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Genome Center

Center code: WIBR

Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L9166
 Center clone name: 795 B 6

----- Summary Statistics

Sequencing vector: M13; M77815; 31% of reads
 Sequencing vector: Plasmid; n/a; 69% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Assembly quality: 210748 bases at least Q40
 Consensus quality: 213655 bases at least Q30
 Consensus quality: 215058 bases at least Q20
 Insert size: 19400; agarose-fp
 Insert size: 216246; sum-of-contigs
 Quality coverage: 11.9 in Q20 bases; agarose-fp
 Quality coverage: 10.7 in Q20.

* NOTE: This is a "working draft" sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 8149: contig of 8149 bp in length
 * 8150 8249: gap of 100 bp
 * 8250 9592: contig of 1343 bp in length
 * 9593 9692: gap of 100 bp
 * 9693 10733: contig of 1041 bp in length
 * 10734 10833: gap of 100 bp
 * 10834 13519: contig of 2666 bp in length
 * 13520 13619: gap of 100 bp
 * 13620 17510: contig of 3891 bp in length
 * 17511 17610: gap of 100 bp
 * 17611 24602: contig of 6992 bp in length
 * 24603 24702: gap of 100 bp
 * 24703 35434: contig of 10732 bp in length
 * 35435 35534: gap of 100 bp
 * 35535 124474: contig of 88940 bp in length
 * 124475 124574: gap of 100 bp
 * 124575 134664: contig of 10090 bp in length
 * 134665 134764: gap of 100 bp
 * 134765 162343: contig of 27579 bp in length
 * 162344 162443: gap of 100 bp
 * 162444 208917: contig of 46674 bp in length
 * 208918 209017: gap of 100 bp
 * 209018 217346: contig of 8329 bp in length.

FEATURES

source

1..217346
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="19"
 /map="19"

1..8149
 /clone="RP11-795B6"
 /clone_1fb="RP11-11 Human Male BAC"

1..8149
 /note="assembly_fragment"

clone_end:SP6
 vector_side:left"

misc_feature
 8250..9592
 /note="assembly_fragment"
 misc_feature
 9693..10733

QY	1801	AGTTAAACAGCTCTCATTTAGCAGGGGTGTGTGTTTCAACAGTATGTTAGGAAGCTGGGGAATT	1860
Db	196613	AGTTAAACAGCTCTCATTTAGCAGGGGTGTGTGTTTCAACAGTATGTTAGGAAGCTGGGGAATT	196554
QY	1861	CAGAGTACTCCAGTCCCATGGCTATGAAAAGCTCCCCCAATTGATCAAACTTGACAA	1920
Db	196553	CAGAGTACTCCAGTCCCATGGCTATGAAAAGCTCCCCCAATTGATCAAACTTGACAA	196494
QY	1921	ATGCAACACCTCCCAAGCTCTCCCATTTCTTCTCTGTGACCCCTGGGTGTGGGGGGGTGGG	1980
Db	196493	ATGCAACACCTCCCAAGCTCTCCCATTTCTTCTGTGACCCCTGGGTGTGGGGGGGTGGG	196434
QY	1981	TTGCGAGGGGGGAAAACCTTTTA 2001	
Db	196433	TTGCGAGGGGGGAAAACCTTTTA 196413	
RESULT 8			
AC130782/c			
LOCUS	AC130782/c		
DEFINITION	Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12		
ACCESSION	AC130782		
VERSION	AC130782.2 GI:25167101		
KEYWORDS	HTG, HTGS PHASE2, HTGS DRAFT.		
SOURCE	Pan troglodytes (chimpanzee)		
ORGANISM	Pan troglodytes		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
AUTHORS	1 (bases 1 to 200792)		
REFERENCE	Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blankley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Carlsang,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P., Lee-Lin,S.-Q., Legeaspi,R., Maduro,Q.L., Maduro,V.B., Marquis,E.H., Masiello,C., Maskeri,B., McDowell,D., Padurigan,C., Pearson,R., Portnoy,M.E., Prasad,A., Padurix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C., Staritrop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.T., Wehrber,K.D., Wiggins,L., Young,A. and Green,E.D.		
TITLE	NISC Comparative Sequencing Initiative		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 200792)		
AUTHORS	Green,E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717		
REFERENCE	Groveont Circle, Gaithersburg, MD 20877, USA		
AUTHORS	3 (bases 1 to 200792)		
TITLE	Green,E.D.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717		
AUTHORS	Groveont Circle, Gaithersburg, MD 20877, USA		
TITLE	On Nov 22, 2002 this sequence version replaced gi:22218452.		
JOURNAL	----- Genome Center		
COMMENT	Center: NIH Intramural Sequencing Center		
	Center code: NISC		
	Web site: http://www.nisc.nih.gov		
	Contact: nisc.zoo@hgti.nih.gov		
	----- Project Information		
	Center project name: dnz		
	Center clone name: 355A20		

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence

contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 196656 bases at least Q40
Consensus quality: 197883 bases at least Q30
Consensus quality: 198879 bases at least Q20
Insert size: 215000; agarose-fp
Insert size: 199692; sum-of-configs
Quality coverage: 9.03% in Q20 bases; agarose-fp
Quality coverage: 9.73% in Q20 bases; sum-of-configs

* NOTE: This is a 'working draft' sequence. It currently consists of 12 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.

- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

*	1	28305:	contig of 28306 bp in length
*	28307	28405:	gap of unknown length
*	28407	37855:	contig of 9450 bp in length
*	37857	37956:	gap of unknown length
*	37957	73522:	contig of 35366 bp in length
*	73523	73622:	gap of unknown length
*	73623	83567:	contig of 9945 bp in length
*	83568	83667:	gap of unknown length
*	83668	88811:	contig of 5150 bp in length
*	88818	88917:	gap of unknown length
*	88918	125511:	contig of 36694 bp in length
*	125512	125711:	gap of unknown length
*	125712	125871:	contig of 34168 bp in length
*	125880	155979:	gap of unknown length
*	155980	174696:	contig of 14719 bp in length
*	174699	174798:	gap of unknown length
*	174799	186382:	contig of 11584 bp in length
*	186383	186482:	gap of unknown length
*	186483	193444:	contig of 6862 bp in length
*	193335	193444:	gap of unknown length
*	193345	193365:	contig of 5919 bp in length
*	193346	199465:	gap of unknown length
*	199364	199465:	contig of 1339 bp in length.
*	199464	199479:	contig of 1339 bp in length.

FEATURES

```
source
1. 200792
/organism="Pan troglodytes"
/mol_type="Genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"
1. 28306
/misc_feature
/note="assembly_fragment
clone end:Sp6
vector side:left"
28407..37856
/note="assembly_fragment"
37957..73522
/note="assembly_fragment"
73623..83567
/note="assembly_fragment"
83668..88817
/note="assembly_fragment"
/note="assembly_fragment"
88918..112611
/note="assembly_fragment"
125712..159879
/note="assembly_fragment"
159980..174688
/note="assembly_fragment"
174799..186382
/note="assembly_fragment"
```

misc_feature 186483..193344
/note="assembly_fragment"
misc_feature 193445..199363
/note="assembly_fragment"
misc_feature 199464..200792
/note="assembly_fragment"
clone end:T7
vector_side:right"

ORIGIN

Query Match 96.3%; Score 1927.2; DB 2; Length 200792;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1975; Conservative 0; Mismatches 23; Indels 4; Gaps 3;
QY 1 GACACCAAGAGCCCCCAAGGTGAGTCCAGGTTCTTCTTGAATACCAATCTCTGC 60
DB 73157 GACAAACCAAGAGCCCCCAAGGTGAGTCCAGGTTCTTCTTGAATACCAATCTCTGC 73098
QY 61 CGCCTTCATCTTCTCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 120
DB 73097 CGCCTTCATCTTCTCACTTCTCATTTGTTCTGTTGACAGTGCATTCCTTAAG 73038
QY 121 TCCCTCAGTGGTGAATATCAGCGTGTAACTAGAAAAGGTGAGAGTCTTACCGA 180
DB 73037 TCCCTCAGTGGTGAATATCAGCGTGTAACTAGAAAAGGTGAGAGTCTTACCGA 72978
QY 181 GACGATAGATGACACCATGTTCTGCGCCGGTGCAAAAGCAGGTAGAGATCTCTGCCAGG 240
DB 72977 GACGATAGATGACACCATGTTCTGCGCCGGTGCAAAAGCAGGTAGAGATCTCTGCCAGG 72918
QY 241 TGAAGACACCTCTCTTATTAGAGATACACATGAGTGCACCTGGTAACTGAGAGC 300
DB 72917 TGAAGACACCTCTCTTATTAGAGATACACATGAGTGCACCTGGTAACTGAGAGC 72858
QY 301 GTTCCAAATTTCTGAGATCCAGCAATTCAGAGACAGTCAAGACCCCTGTTCTACAGA 360
DB 72857 GTTCCAAATTTCTGAGATCCAGCAATTCAGAGACAGTCAAGACCCCTGTTCTACAGA 72798
QY 361 GCTCATACCTTAGAGTGTGTTTATGATAAATATGCTGAGCTGTTATGTCATTTTC 420
DB 72797 GCTCATACCTTAGAGTGTGTTTATGATAAATATGCTGAGCTGTTATGTCATTTTC 72738
QY 421 CAGTTTTTATGATCCACATTTAAACAGGTAAAGGCTGGGCGAGTGGCTCACACT 480
DB 72737 CAGTTTTTATGATCCACATTTAAACAGGTAAAGGCTGGGCGAGTGGCTCACACT 72678
QY 481 GTATTCACAGACTTTTGGAGGCTGAGGACAGATCACTTTTGTCAAGAGTTTGAGA 540
DB 72677 GTATTCACAGACTTTTGGAGGCTGAGGACAGATCACTTTTGTCAAGAGTTTGAGA 72618
QY 541 CTAGCTGGCCCAACATGGCCAACTCTGTCTCTTAAAAAAATACAAATTTAGCTGGC 600
DB 72617 CTAGCTGGCCCAACATGGCCAACTCTGTCTCTTAAAAAAATACAAATTTAGCTGGC 72558
QY 601 ATGGTGGCGGCGCTGTAAATCTCAGCTGCTCAGAGGCGGAGACACAGAAATCACTTAA 660
DB 72557 ATGGTGGCGGCGCTGTAAATCTCAGCTGCTCAGAGGCGGAGACACAGAAATCACTTAA 72498
QY 661 ACCCAGAGAGTGAAGTTCAGTGAAGTGAATGCTGCACT -CACTTCAACTGGAGAGA 719
DB 72497 ACCCAGAGAGTGAAGTTCAGTGAAGTGAATGCTGCACTGCACTCAACTGGAGAGA 72438
QY 720 CAGAGTACACTTTTGTCTCAAAAAAGAAAAAACAGTAAAGAAACAGGTGAA 779
DB 72437 CAGAGTACACTTTTGTCTCAAAAAAG -AAAAAACAAGTAAAAAAGAAACAGGTGAA 72379
QY 780 GTTAACCTTAATACCAATGATCCCAATATCAATCTTCAAGTGTAAATTAATATAA 839
DB 72378 GTTAACCTTAATACCAATGATCCCAATATCAATCTTCAAGTGTAAATTAATATAA 72319
QY 840 AACAAATATGATAGATTAATCTTCAATCTTCTGTTTCAATTAATAGCTTTGAAAG 899
DB 72318 AACAAATATGATAGATTAATCTTCAATCTTCTGTTTCAATTAATAGCTTTGAAAG 72259

QY 900 TGAATATATATGTTATGCTGACAGCATCTCAATTTGACCTAGCTACATTTGAGGTCT 959
DB 72258 TGAATATATATTTATGCTGACAGCATCTCAATTTGACCTAGCTACATTTGAGGTCT 72199
QY 960 CAGTAGCCATATGCTGCTAGCACTTACTGATTTGATGACCGGATCTAGAGGAAAGAT 1019
DB 72198 CAGTAGCCATATGCTGCTAGCACTTACTGATTTGATGACCGGATCTAGAGGAAAGAT 72139
QY 1020 CAGGCTGTTTGTATGTTGGCAGGTTGTGCACTGCATTAAGATACATATCTAATAG 1079
DB 72138 CAGGCTGTTTGTATGTTGGCAGGTTGTGCACTGCATTAAGATACATATCTAATAG 72079
QY 1080 GGGCACTCCGTTTACAGATGATGATTTTGGAGTTTTCAGGGGTGTGATTAAGTGT 1139
DB 72078 GGGCACTCCAGTTTACAGATGATGATTTTGGAGTTTTCAGGGGTGTGATTAAGTGT 72019
QY 1140 CTGTGTTCAACAAATCTGTAATATGACAGTTTCTAGCAAGTCTGTAAAAATCTTTC 1199
DB 72018 CTGTGTTCAACAAATCTGTAATATGACAGTTTCTAGCAAGTCTGTAAAAATCTTTC 71959
QY 1200 AGAAGGAAAAAGAAATCTGTAGGTATTTTACAAAGATATTTAATACAGGGAAT 1259
DB 71958 AGAAGGAAAAAGAAATCTGTAGGTATTTTACAAAGATATTTAATACAGGGAAT 71899
QY 1260 AATTGCAAGCTCTGGAAGGCTGAGAGAACAAAGTTAAATTAATTAATTAATCTGTG 1319
DB 71898 AATTGCAAGCTCTGGAAGGCTGAGAGAACAAAGTTAAATTAATTAATTAATCTGTG 71839
QY 1320 CAAGAAATCTGCATTAATATAGGCAATTTACAGAGAGGTAAAGTTAAATCCCAATATAA 1379
DB 71838 CAAGAAATCTGCATTAATATAGGCAATTTACAGAGAGGTAAAGTTAAATCCCAATATAA 71779
QY 1380 CATGTTTATGATAGTAAACAAATTAAGGCAATTTCAAAAAAGTGTCTAGGGAGCT 1439
DB 71778 CATGTTTATGATAGTAAACAAATTAAGGCAATTTCAAAAAAGTGTCTAGGGAGCT 71719
QY 1440 CTTTGAAGAGTGGCAATTTGAGAGAGATGATGACAAAGAGTAAACTCGTAG 1499
DB 71718 CTTTGAAGAGTGGCAATTTGAGAGAGATGATGACAAAGAGTAAACTCGTAG 71659
QY 1500 TTTAAGGGGAAAAAGGACCTGCAAAAGCCCTGAGGAGTAAAGAAATTTGGCTGATT 1559
DB 71658 TTTAAGGGGAAAAAGGACCTGCAAAAGCCCTGAGGAGTAAAGAAATTTGGCTGATT 71599
QY 1560 CAAAGAAAGAGGAAACCAATGCAACTGAGAAACAAAGTGGGGGCAACATAGAAAGT 1619
DB 71598 CAAAGAAAGAGGAAACCAATGCAACTGAGAAACAAAGTGGGGGCAACATAGAAAGT 71539
QY 1620 GAGGCTGAGAGTGTAGCAGGGGCGAATGCTCTGCAAGTATTTCTTGTCTCAACACAG 1679
DB 71538 GAGGCTGAGAGTGTAGCAGGGGCGAATGCTCTGCAAGTATTTCTTGTCTCAACACAG 71479
QY 1680 AGCTTCCCTATGTTCTTAATGGAAGCTGATCTGTTGAGAAACAGAAATTTAAATCAAA 1739
DB 71478 AGCTTCCCTATGTTCTTAATGGAAGCTGATCTTAATGGAAGAAACAGAAATTTAAATCAAA 71419
QY 1740 CTGTATCATCAACACAGACCTTCTCTGTAATTCAGGCTCCCAAGGATCTTGAAGACGT 1799
DB 71418 CTGTATCATCAACACAGACCTTCTCTGTAATTCAGGCTCCCAAGGATCTTGAAGACGT 71359
QY 1800 AAGTTTAAACAAGCTCTCATTAAGCAGGGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAT 1859
DB 71358 AAGTTTAAACAAGCTCTCATTAAGCAGGGGTGTGTTTCAACAGTATTAAGAAAGCTGGGAT 71299
QY 1860 TCAGAGTACTCCAGTCCCATGCTATGAAAAGCTCCCAATTTGTAACAACTGACACA 1919
DB 71298 TCAGAGTACTCCAGTCCCATGCTATGAAAAGCTCCCAATTTGTAACAACTGACACA 71239
QY 1920 AATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGCTCTGGGTGTGGGGGGTGG 1979
DB 71238 AATGCAACACTCCCAAGCTCTCCCAATTTCTCTGTGCTCTGGGTGTGGGGGGTGG -GG 71181

repeat_region	9528. . 9727
repeat_region	/rpt_family="MER1_type"
repeat_region	12230. .12519
repeat_region	/rpt_family="A1u"
repeat_region	13130. .13267
repeat_region	/rpt_family="MER2_type"
repeat_region	13268. .13806
repeat_region	/rpt_family="MER1_type"
repeat_region	13807. .14000
repeat_region	/rpt_family="MER2_type"
repeat_region	14571. .14710
repeat_region	/rpt_family="MIR"
repeat_region	14747. .14772
repeat_region	/rpt_family="(TA)n"
repeat_region	14775. .15038
repeat_region	/rpt_family="A1u"
repeat_region	15043. .15760
repeat_region	/rpt_family="ERV1"
repeat_region	15838. .15913
repeat_region	/rpt_family="L1"
repeat_region	16094. .16263
repeat_region	/rpt_family="ERV1"
repeat_region	16266. .16328
repeat_region	/rpt_family="MER1_type"
repeat_region	16328. .16352
repeat_region	/rpt_family="(AAAA)n"
repeat_region	16370. .16583
repeat_region	/rpt_family="MALR"
repeat_region	17083. .17301
repeat_region	/rpt_family="MER1_type"
repeat_region	17470. .17565
repeat_region	/rpt_family="L2"
repeat_region	17715. .18240
repeat_region	/rpt_family="MALR"
repeat_region	18302. .18341
repeat_region	/rpt_family="MALR"
repeat_region	18421. .18816
repeat_region	/rpt_family="L1"
repeat_region	18980. .19106
repeat_region	/rpt_family="A1u"
repeat_region	19141. .19442
repeat_region	/rpt_family="A1u"
repeat_region	21116. .21322
repeat_region	/rpt_family="(TTC)n"
repeat_region	21328. .23189
repeat_region	/rpt_family="L1"
repeat_region	23215. .23315
repeat_region	/rpt_family="U6"
repeat_region	23316. .23365
repeat_region	/rpt_family="L1"
repeat_region	23641. .23897
repeat_region	/rpt_family="MIR"
repeat_region	23934. .24124
repeat_region	/rpt_family="MIR"
repeat_region	25057. .25137
repeat_region	/rpt_family="GA-rich"
repeat_region	25392. .25488
repeat_region	/rpt_family="7SK"
repeat_region	25497. .26487
repeat_region	/rpt_family="L1"
repeat_region	26514. .26581

	Query Match	Best Local Similarity	14.3%	Score 286.4	DB 9	length 142003
	Matches 422	Conservative 0	Mismatches 151	Indels 12	Gaps 4	
QY	416	ATTTCACGTTTTTAGTACCACTTAAAAACAGTAAAAAGGCTGGGGCGCATGGCTCA				475
Db	27610	ACTTAATAATATAATAGTACCACTTAAAAAA--TAAAGGCTGGGGCGCATGGCTCA				27666
QY	476	CACCTGTAAATCCAGACACTTTGGAGAGCTGAGGACAGATCACTTTGGTCAGAGATT				535
Db	27667	CACCTGTAAATCCAGACACTTTGGAGAGCTGAGGACAGTGAATCACTCAAGATCAGAGATT				27726

QY	536	TGAGCATAGGCTGGCCCAATGGCGAACTGTCTCTAAAAAATAACAAATATAC	595
Db	27727	CGAACACAGGCTGACCAACATGGCGAAACCC---CTCTCTACTAAAAATATATAATTTAT	27785
QY	596	CTGGCATGTGTGGCGGGCGCTGTAACTCAGCTGTCAGAGGCGCCGAGACACAAGATCA	655
Db	27784	CCGGCGTGTGTGTGGGCTGTAACTCTAACTACTTGGAGGCTGTAGACAGAGATCA	27843
QY	656	CTTAAACCCGAGAGGTGTGAGGTGTGACGTGAGCTGAGATTTGTGCCACT-CACCTCAACCTG	714
Db	27844	CTTGAACCCGAGAGGTGTGAGGTGTGACGTGAGCGAGATGTGTCAATTGTTCAGCCTG	27903
QY	715	GGAGACAGAGTGCACACTTTTGTCTCAAAAAGAAAAAACAAGTAAAAAAGAAACAG	774
Db	27904	GGTACGCTATGTGAGCTCCGCTCABAAAAAAGTAAAAAAGTAAAAATATTTT	27963
QY	775	GTCGAATTAACTTTAATATACCAAGTATCCCAATACATCATTTCAAAAGTGTATTTA	834
Db	27964	TAAATAAATATTTTATTTATTCATTTATAGCAAAATATGATCATTTTCAACATGAATCA	28023
QY	835	TATTAACCATTTATGATGATGACTTTTCACTTTTCTGTGTTTCATATTAAGCTTT	894
Db	28024	TATTTAAAACTTTATCAAGATTAATTTTACATTTCTTT-----TTTTCATACAAATTTCTTA	28078
QY	895	GAAAGTAGATATATATGTATATGCTGACAGACATCTCAATTTTGAAGTAACTATTCAG	954
Db	28079	GAAATCTGTCTGCATTTATATTTTCAAGGACATTTCAATTTGAGCTACCACTATTTTCA	28138
QY	955	GTCGTCAATGACCAATGTGGCTAGACAGTTACGTATTTGAGTGGACGGATC	1006
Db	28139	GTTGTTATTTATGACACGTGGCTAATGGCTGCATTTAGATTAAGCTCACAGCT	28190

LOCUS	AC133961	177877 bp	DNA	linear	PRI 03-JAN-2002
DEFINITION	Homo sapiens BAC clone RP13-494C23 from 4, complete sequence.				
ACCESSION	AC133961				
VERSION	AC133961.3 GI:26190601				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	1 (bases 1 to 17787)				
JOURNAL	Sulston, J.E. and Waterston, R.				
GENOME	Toward a complete human genome sequence				
REVIEW	Genome Res. 8 (11), 1097-1108 (1998)				
PUBMED	99063792				
REFERENCE	2 (bases 1 to 17787)				
AUTHORS	Isak, A., Bielicki, L., Creason, K. and Cotton, M.				
TITLE	The sequence of Homo sapiens BAC clone RP13-494C23				
JOURNAL	Unpublished (2001)				
REFERENCE	3 (bases 1 to 17787)				
AUTHORS	Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-SEP-2002) Genome Sequencing Center, Washington				
UNIVERSITY	University School of Medicine, 4444 Forest Park Parkway, St. Louis,				
MO	63108, USA				
REFERENCE	4 (bases 1 to 17787)				
AUTHORS	Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-NOV-2002) Genome Sequencing Center, Washington				
UNIVERSITY	University School of Medicine, 4444 Forest Park Parkway, St. Louis,				
MO	63108, USA				
REFERENCE	5 (bases 1 to 17787)				
AUTHORS	Waterston, R.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-DEC-2002) Genome Sequencing Center, Washington				
UNIVERSITY	University School of Medicine, 4444 Forest Park Parkway, St. Louis,				
MO	63108, USA				

REFERENCE 6 (bases 1 to 177787)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Dec 7, 2002 this sequence version replaced gi:24580489.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu
----- Summary Statistics
Center project name: H_FH0494C23

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCT-13 Human Female BAC Library was constructed using improved cloning techniques developed by Kazutoyo Osoegawa. The library was generated by Baohui Zhao in our laboratory. Construction was funded by a grant from the National Human Genome Research Institute (NHGRI, NIH) (#1R01RG0165-03). The library was generated according to the new NHGRI/DOE 'Guidance on Human Subjects in Large-Scale DNA Sequencing'.
Female blood was obtained via a double-blind selection protocol. Female blood DNA was isolated from one randomly chosen donor (out of 10 female donors) and partially digested with a combination of EcoRI and EcoRI Methylase for library segments 1k2 or either MboI or PvuII for library segments 3k4. Size selected DNA was cloned into the pBACe3.6 vector between the EcoRI sites for library segments 1k2 or the BamHI sites for library segments 3k4. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies). The library has been arrayed into 384-well microtiter dishes and also gridded onto 22x22cm nylon high density filters for screening by probe hybridization.

NEIGHBORING SEQUENCE INFORMATION:
The sequence from base 67288 to base 67340 was derived from one plasmid subclone.

Polymorphisms have been identified between AC020706 and this sequence.

Data from AC020706 was used to finish this clone.

This sequence is not the entire insert of the clone. This clone is overlapped by AC020706 and AC093660.

FEATURES

source

1.177787
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP13-494C23"

/clone_11b="RPCT-13"
4..70
/rpt_family="MIR"
703..833
/rpt_family="Alu"
852..875
/rpt_family="AT_rich"
911..1138
/rpt_family="Alu"
1137..1191
/rpt_family="(CAAAA)n"
1341..1408
/rpt_family="GA-rich"
1432..1577
/rpt_family="GA-rich"
2115..2458
/rpt_family="12"
2562..2603
/rpt_family="MER1_type"
2773..2868
/rpt_family="MER1_type"
2932..3240
/rpt_family="Alu"
3264..3462
/rpt_family="MALR"
3463..3760
/rpt_family="Alu"
3761..3934
/rpt_family="MALR"
5508..5796
/rpt_family="Alu"
6977..7240
/rpt_family="Alu"
7250..7270
/rpt_family="AT_rich"
7524..7661
/rpt_family="MIR"
7891..8423
/rpt_family="ERV1"
9624..9825
/rpt_family="MIR"
10105..10163
/rpt_family="12"
10414..10721
/rpt_family="Alu"
11332..11671
/rpt_family="11"
12367..12389
/rpt_family="AT_rich"
13174..13474
/rpt_family="Alu"
13507..13850
/rpt_family="Alu"
13834..13969
/rpt_family="Alu"
14004..14059
/rpt_family="polypyrimidine"
16122..16152
/rpt_family="(T)n"
16824..17116
/rpt_family="Alu"
17122..17421
/rpt_family="Alu"
17422..18113
/rpt_family="11"
18681..18727
/rpt_family="12"
20147..20162
/rpt_family="MIR"
20165..20386
/rpt_family="MER1_type"
20387..20502
/rpt_family="MIR"

	repeat_region	20503	.20614	/rpf_family="TA)n"
	repeat_region	20617	.20782	/rpf_family="WIR"
	repeat_region	21217	.21523	/rpf_family="L2"
	repeat_region	21726	.21824	/rpf_family="CR1"
	repeat_region	21947	.22277	/rpf_family="MER1_type"
	repeat_region	24511	.24634	/rpf_family="L2"
	repeat_region	24922	.25153	/rpf_family="MER1_type"
	repeat_region	25154	.25413	/rpf_family="Alu"
	repeat_region	25414	.25462	
Query Match	13.8%	Score 276.8	DB 9	Length 17787
Best Local Similarity	68.9%	Pred. No. 6.2e-55		
Matches 460	Conservative	0	Mismatches 187	Indels 21
			Gaps	5
Qy	311	TCTGAGATCCAGCAATTTGCCAAGACAGTCAAGACCCCTGTTCTCAAGAGCTCATACC	370	
Db	58503	TCTGGCATTTTCTTTTCCCAAGAAAGTCACTGACACTCAATTGAAATGACACTATC	58444	
Qy	371	TAGAGTAGTGTTTATAGTAAATTAATGCTGACCTGCTTATGTCATTTCCAGTTTCTTA	430	
Db	58443	CAGTTGAAATACATATGTAACCAAAATTTCAAGCCACATATATCATTTTAAATTCCTTA	58384	
Qy	431	GTAACCATTTAAACAGGTAAATAAGCTGGGCGAGTGGCTCACACCTGTAAATCCGAG	490	
Db	58383	GTACTATGTTTAAATAAGTAAAGAGGCGCGGGGCGAGTGGCTCACACCTGTATCTTAG	58324	
Qy	491	CACTTTGGGAGGCTGAGGCAAGCATCACTTTGTGAGAGATTGAGACTAGCTGAGC	550	
Db	58323	CACTTTGGGAGGCGGAGGAGGAGTGAATCACTGAGTCAAGAGATTGAGACCAAGCTGAGC	58264	
Qy	551	CAACATGGCGAATCTGTCTCTTAAAAAAATTAAGCTGCGATGAGTGGCGG	610	
Db	58263	CAACATGGCAAAACCTGTCTCT-ACGTAAATACAAATAATAGCTGGGCAATGTGGCAC	58205	
Qy	611	GGCCTGTGAATCTCAAGCTGCTCAGAGGCGGAGACAAAGATTCATTAAACCCAGAGG	670	
Db	58204	ATGCTGTGATCTAGCTGCTCAGAGGCTGAGGCAAGAAATCGCTAGAACCCGGGAGG	58145	
Qy	671	TGAGAGTTGACGAGGAGTGTGTCCT-CACTCCAACTGGGAGCAGAGTGACA	729	
Db	58144	CGAGAGTTGACATAGCGAGATGGCCCTATTGACATCCAGCTGGGCGCAAGAGGACA	58085	
Qy	730	CTTTTGTCTCAAAAGAAAAAACAAGTAAATAAGAAAGAAAGAGTGAATTAATCTTA	789	
Db	58084	CTTCATCTCAAAAAAAGAAAAAAGAAAAAAGTAAAGAGAGGATTAATTA	58025	
Qy	790	ATAACCCAGTAT-----CCCAATACATCATTTCAAAGTGAATTAATA	836	
Db	58024	TTAATGTATTTTATTTAACTCAAGTATCAAAATATGTATGATCAACATCAATTAATA	57965	
Qy	837	TAAACAATTTGAATGAGTACTTTCATTTCTTTTCTTAATTAATTAAGCTTTGA	896	
Db	57964	TAAA-ATTATTAATAGATATTTTTCAGCC-----CTTTTATTAACAAAGGTTTGA	57911	
Qy	897	AAGTAGATATATGTTATGCTGACAGACATCTCAATTTGAGTACTAGCTAATTCAGGT	956	
Db	57910	AATTCATGATATTTTACCTTACAGAGATCTCAATCAAGACAGCCACATTACAAGT	57851	
Qy	957	GCTCAGTA 964		
Db	57850	GCTTAATA 57843		

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Human7.7LH19	138411 bp DNA linear PRI 11-AUG-2000	AL121958	AL121958.6	GI:8247268	HTG; PRIM2A.	Homo sapiens (human)	1 (bases 1 to 138411)	Brown, J.	Direct Submision	Submitted (27-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk	<p>Human7.7LH19 is a DNA sequence from clone RPI-7LH19 on chromosome 6 Contigs part of the PRIM2A gene encoding DNA primase large subunit (p58), ESTs, STS and GSSs, complete sequence.</p> <p>This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unseq' feature key.</p> <p>The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/NGP/Ch6</p> <p>RPI-7LH19 is from the library RPI-1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/</p> <p>VECTOR: pCYPAC2</p> <p>This sequence is the entire insert of clone RPI-7LH19 The true left end of clone RPI-216J23 is at 87497 in this sequence. The true right end of clone RPI-422B11 is at 29670 in this sequence.</p>
FEATURES	<p>Location/Qualifiers</p> <p>1..138411</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="6"</p> <p>/clone="RPI-7LH19"</p> <p>/clone_1lb="RPI-1"</p> <p>/complement(477..1095)</p> <p>/note="match: GSS: Em:AQ0506907"</p> <p>/complement(609..1046)</p> <p>/note="match: GSS: Em:AQ437147"</p> <p>/complement(620..1089)</p> <p>/note="match: GSS: Em:AQ819330"</p> <p>/complement(621..1083)</p> <p>/note="match: GSS: Em:AQ437147"</p> <p>/complement(761..1000)</p> <p>/note="match: GSS: Em:BT1825"</p> <p>/complement(977..1090)</p> <p>/note="match: GSS: Em:AQ55044"</p> <p>1514..1858</p> <p>/note="match: GSS: Em:AQ176950"</p> <p>1545..1851</p> <p>/note="match: GSS: Em:AQ176950"</p> <p>1857..2161</p> <p>/note="AluX repeat: matches 1..309 of consensus"</p> <p>2510..8661</p>										
SOURCE											

```
misc_feature /note="L1PA4 repeat: matches 13. .6142 of consensus"  
8448. .8976  
/note="match: GSS: Em:AQ347835"  
misc_feature 8662. .8847  
/note="match: GSS: Em:AQ347835"  
8848. .9156  
repeat_region /note="MLT11 repeat: matches 40. .410 of consensus"  
9172. .9481  
/note="AluSq repeat: matches 5. .313 of consensus"  
9692. .9804  
repeat_region /note="MLT11 repeat: matches 267. .371 of consensus"  
9824. .10043  
repeat_region /note="MIR repeat: matches 18. .251 of consensus"  
10338. .10689  
/note="MLT2FA repeat: matches 9. .356 of consensus"  
complement(10977. .11390)  
misc_feature /note="match: SRS: Em:G42414"  
complement(11825. .12306)  
/note="match: GSS: Em:AQ141424"  
12402. .12825  
/note="match: GSS: Em:AQ266412"  
12647. .12722  
repeat_region /note="MIR repeat: matches 1. .79 of consensus"  
13978. .14287  
repeat_region /note="AluDb repeat: matches 1. .310 of consensus"  
14664. .15035  
/note="L2 repeat: matches 2294. .2673 of consensus"  
15036. .15342  
repeat_region /note="AluSp repeat: matches 1. .308 of consensus"  
15343. .15421  
/note="L2 repeat: matches 2673. .2750 of consensus"  
15485. .15534  
repeat_region /note="MIR repeat: matches 213. .262 of consensus"  
15581. .15703  
repeat_region /note="MIR repeat: matches 110. .234 of consensus"  
16310. .16535  
/note="MIR repeat: matches 21. .231 of consensus"  
16597. .17598  
repeat_region /note="L1PA7 repeat: matches 5135. .6145 of consensus"  
17617. .17827  
repeat_region /note="AluSp repeat: matches 83. .309 of consensus"  
17828. .19349  
/note="L1PA7 repeat: matches 3624. .5146 of consensus"  
complement(19723. .20194)  
misc_feature /note="match: GSS: Em:AQ016849"  
20519. .20665  
/note="match: GSS: Em:AQ371979"  
20807. .21136  
misc_feature /note="match: GSS: Em:AQ508843"  
20807. .21096  
misc_feature /note="match: GSS: Em:AQ508843"  
20818. .21272  
/note="match: GSS: Em:AQ209413"  
20818. .21096  
misc_feature /note="match: GSS: Em:AQ209413"  
20831. .21139  
misc_feature /note="match: GSS: Em:AQ514718"  
20902. .21139  
misc_feature /note="match: GSS: Em:AQ093115"  
21097. .21176  
/note="40 copies 2 mer ct 76% conserved"  
21524. .22065  
/note="match: GSS: Em:AQ50328"  
22856. .23094  
repeat_region /note="AluSq repeat: matches 1. .239 of consensus"  
23102. .23135  
repeat_region /note="L17 copies 2 mer aa 88% conserved"  
24117. .24204  
/note="MER5B repeat: matches 1. .90 of consensus"  
24883. .25185  
repeat_region /note="AluY repeat: matches 1. .300 of consensus"  
25364. .25449  
/note="MSTD repeat: matches 1. .96 of consensus"
```

```
repeat_region 25932. .26300  
/note="L1MB8 repeat: matches 5779. .6162 of consensus"  
repeat_region 26483. .26600  
/note="L2 repeat: matches 2583. .2703 of consensus"  
repeat_region 27463. .27633  
/note="AluSp repeat: matches 137. .308 of consensus"  
misc_feature complement(27586. .28035)  
/note="match: GSS: Em:AQ753792"  
complement(27650. .27736)  
/note="match: GSS: Em:AQ753792"  
27737. .28017  
repeat_region /note="AluDb repeat: matches 1. .276 of consensus"  
28766. .28860  
repeat_region /note="MIR repeat: matches 85. .188 of consensus"  
29043. .29225  
/note="MIR repeat: matches 0. .174 of consensus"  
29387. .29497  
repeat_region /note="PRAM repeat: matches 29. .143 of consensus"  
29585. .29879  
repeat_region /note="AluSp repeat: matches 1. .296 of consensus"  
31001. .31343  
repeat_region /note="L1MD2 repeat: matches 5957. .6332 of consensus"  
31344. .31643  
/note="AluSc repeat: matches 1. .301 of consensus"  
31644. .31712  
repeat_region /note="L1MD2 repeat: matches 5890. .5957 of consensus"  
31714. .31863  
repeat_region /note="L1MD3 repeat: matches 7589. .7735 of consensus"  
31874. .32598  
/note="L1MD2 repeat: matches 5386. .6122 of consensus"  
33641. .33934  
repeat_region /note="AluXk repeat: matches 3. .298 of consensus"  
33933. .33941  
/note="single clone region"  
34941. .35324  
misc_feature /note="match: GSS: Em:B55255"  
35189. .35758  
repeat_region /note="L2 repeat: matches 2189. .2747 of consensus"  
complement(36539. .37046)  
misc_feature /note="match: GSS: Em:AQ558962"  
36809. .37468  
/note="match: GSS: Em:AQ541391"  
36811. .37282  
misc_feature /note="match: GSS: Em:AQ421653"  
36831. .37467  
/note="match: GSS: Em:AQ547524"  
36992. .37051  
repeat_region /note="MER85 repeat: matches 86. .139 of consensus"  
37054. .37177  
repeat_region /note="MER85 repeat: matches 1. .139 of consensus"  
37956. .38473  
misc_feature /note="match: GSS: Em:AQ246505"  
complement(37968. .38506)  
/note="match: GSS: Em:AQ427574"  
misc_feature  
  
Query Match 13.7%! Score 274.4; DB 9; Length 138411;  
Best local Similarity 78.0%; Pred. No. 2.3e-54;  
Matches 415; Conservative 0; Mismatches 91; Indels 26; Gaps 6;  
  
QY 472 CTCACACCTGTATATCCAGACACTTTGGAGGCTGAGGCGAGGATACACTTTGGTCAGG 531  
|||||  
Db 43069 CTCACACCTGTATATCTCAGACACTTTGGAGGCGGCGAGATCACTGAGGTCAGG 43128  
|||||  
QY 532 AATTGAGACTGACCTGGCCCAACATGCGGAATCTGTCTCTCAAAAAAATACAAAAT 591  
|||||  
Db 43139 AATTGAGACTGACCTGGCCCAACATGGAACCTTATCTCT-ACATAAAATACAAAAT 43187  
|||||  
QY 592 TAGCTGCGATGTCGCGCGCGCTGTATATCTCAGCTGTCAGAGCGGAGACACAAGA 651  
|||||  
Db 43188 TAGCAGGCGATGTCGTGGGTCCTGTATATCCACTACTCAGAGGCTGAGGCGAGAGA 43247  
|||||  
QY 652 ATCACTTAACCCAGAGAGGTGAGGTTGCACTGAGACTGATGTGCCACT-CACTCCAA 710  
|||||
```


Db	98545	GAGGTCAGGAGTTGAGACCAAGCCTGGTCAACAATGGTGAACCCCGCTCT-ACATAAAA	98487
Qy	583	TACAAAAATTAGCTGGCATGTGTGGCGGCGCTGTATCTCAGCTGTCAGAGGCCGA	642
Db	98486	CACAAAAATTAGCGGGGTGTGGTGGACACGCGCTGTATGTTCTCAGTACTACAGGAGCTGA	98427
Qy	643	GACACAAGAAATCATTTAAACCCGAGAGTGGAGTGGCAGTGAAGCTGACATCGTGCCT	702
Db	98426	GGCAGAGGAATCATCTTAAACCTGGGAAACAGAGGTGGCAATTAAGCCAAAGTTGAGCACT	98367
Qy	703	-CACTCCCACTCGGGAACAGATGACACTTTTGTCTCAAAAAAGAAAAAAACAAGT	761
Db	98366	GAACTCCCGCTGGGTGACACAGAGCAAACTCCGCTCCAAAAAAAGTAAAGTAATTTT	98307
Qy	762	AAAAAAGAAACAGGTGAAGTTAACTTTATATACCAATGTATCCAAATPACATCTTC	821
Db	98306	AACGA-----TATATTTATTTAGACTCAATATATCCAAATATATATCATTTTC	98261
Qy	822	AAAGTGTATTAATATATAAACAATATAGATGATATCTTATCATTTCTTTCTGTTTC	881
Db	98260	AATATGTATATCAATATATATAAATGTATTAATGAGGTGTGACACCTTT-----TTTTC	98207
Qy	882	ATATTAAGTCTTTAAAGTGTATATATATGTATATGTATGCTGACAGACAACTCAATTTGGACT	941
Db	98206	ATACTAAGTGTCCCAATCTTAATATGATTTTATCATTTTGACAGACACTTCAATTTAGACT	98147
Qy	942	AGTACATATTTGAGGTGTCAATGACCAATGTGGCTAGACATTTACTGTATTTGGATGGCAC	1001
Db	98146	AGTACATATTTTACTGCTCAAGACCAATGTGGCTGTGGTACTGATGCTACTGTCCAGCAC	98087
Qy	1002	GGATCTAGAGGG	1013
Db	98086	AGCTCTAATCTGG	98075
RESULT 13	AC025163	AC025163	177876 bp DNA linear HTG 26-UN-2001
DEFINITION			Homo sapiens chromosome 3 clone RP11-533F2, WORKING DRAFT SEQUENCE, 17 unordered pieces.
ACCESSION			AC025163
KEYWORDS			HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE			1 (bases 1 to 177876)
AUTHORS			Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C., Albriookes, S.L., Amaralunga, H.C., Aze, J.R., Banks, T., Barbarella, J., Benton, J., Blmage, K., Blankenburg, K., Bonini, D., Bouck, J., Bowie, S., Brevet, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Day, A., L., Dederich, D.A., Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H., Douhaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Geo, D., Garcia, A., Garner, T., Garcia, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haxlar, P., Hayes, A., Hernandez, J., Hernandez, O., Hodson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvan, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichner, O., Lieu, C., Liu, J., Liu, W., Louisleg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Meeshwar, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, Z., Miner, Z., Mitchell, T., Mohabbat, K.,

Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenko, S.,
Oguth, M., Okunoye, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,
Pearty, J., Perez, J., Peters, L., Pickens, R., Primus, E., Pu, L.,
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuboan, I., Rolfe, M.,
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshitari, N.,
Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H.,
Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
Thomas, S., Uman, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
Washington, S., Williams, G., Williamson, A., Wleczyk, R.,
Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission
2 (bases 1 to 17876)
Unpublished
Worley, K.C.

Direct Submission
Submitted (07-Mar-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 25, 2001 this sequence version replaced gi:12831249.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----
Center project name: HAMH
Center clone name: RP11-533F2
----- Summary Statistics -----
Sequencing vector: M13, L08821
Chemistry: Dye-Primer Bodyfly: 1% of reads
Chemistry: Dye-terminator Big Dye: 99% of reads
Assembly program: Phrap: version 0.99039
Consensus quality: 169802 bases at least Q40
Consensus quality: 180431 bases at least Q30
Consensus quality: 185136 bases at least Q20
Estimated insert size: 177736, sum-of-contigs estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 4.2x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 42878: contig of 42878 bp in length
* 42879 42978: gap of unknown length
* 42979 62018: contig of 19040 bp in length
* 62019 62118: gap of unknown length
* 62119 80603: contig of 18485 bp in length
* 80604 80703: gap of unknown length
* 80704 98288: contig of 17585 bp in length
* 98289 98388: gap of unknown length
* 98389 113934: contig of 15546 bp in length
* 113935 114035: gap of unknown length
* 114035 125339: contig of 11305 bp in length
* 125340 125439: gap of unknown length
* 125440 138329: contig of 12890 bp in length
* 138330 138429: gap of unknown length
* 138430 145577: contig of 7148 bp in length
* 145578 145677: gap of unknown length
* 145678 150952: contig of 5275 bp in length
* 150953 151052: gap of unknown length
* 151053 153923: contig of 4871 bp in length
* 153924 155924: gap of unknown length


```
repeat_region      4101..4399
                    /note="AluXs repeat: matches 1..299 of consensus"
repeat_region      4400..4765
                    /note="L1 repeat: matches 4004..4439 of consensus"
repeat_region      4772..5034
                    /note="L1M4 repeat: matches 3188..3469 of consensus"
repeat_region      5035..5355
                    /note="AluJo repeat: matches 1..309 of consensus"
misc_feature        complement(5142..5568)
                    /note="match: GSS: Em:B71585"
misc_feature        complement(5243..5586)
                    /note="match: GSS: Em:B71584"
repeat_region      5356..5395
                    /note="L1M4 repeat: matches 3151..3188 of consensus"
repeat_region      5942..6441
                    /note="L2 repeat: matches 1484..2055 of consensus"
repeat_region      7093..7168
                    /note="MIR repeat: matches 97..174 of consensus"
repeat_region      7381..7748
                    /note="MER47A repeat: matches 1..365 of consensus"
repeat_region      8446..8516
                    /note="MIR repeat: matches 68..141 of consensus"
repeat_region      9113..9216
                    /note="26 copies 4 mer agag 66% conserved"
repeat_region      9146..9217
                    /note="36 copies 2 mer ga 77% conserved"
repeat_region      9428..9789
                    /note="L1M4C repeat: matches 1169..1539 of consensus"
repeat_region      11328..11428
                    /note="MER5A repeat: matches 9..112 of consensus"
repeat_region      11564..11680
                    /note="L2 repeat: matches 2581..2710 of consensus"
repeat_region      12030..12247
                    /note="MER21B repeat: matches 414..627 of consensus"
repeat_region      12569..12622
                    /note="MUT2B repeat: matches 394..448 of consensus"
repeat_region      12623..12658
                    /note="18 copies 2 mer ta 100% conserved"
repeat_region      12727..13139
                    /note="MUT2B repeat: matches 1..404 of consensus"
repeat_region      13140..18697
                    /note="HERVL repeat: matches 1..5654 of consensus"
repeat_region      18698..19189
                    /note="MUT2B repeat: matches 1..444 of consensus"
repeat_region      19191..19821
                    /note="MER21B repeat: matches 146..785 of consensus"
repeat_region      20231..20491
                    /note="MUT1A1 repeat: matches 105..365 of consensus"
repeat_region      21266..21603
                    /note="MUT1A1 repeat: matches 3..364 of consensus"
repeat_region      21917..22261
                    /note="MER1B repeat: matches 1..337 of consensus"
Gene                complement(22369..44784)
                    /gene="d4453H5.1"
                    /product="d4453H5.1 (novel protein)"
mRNA                /note="match: CDNA8: Em:AF132953
                    match: EST8: Em:AW702079 Em:AA509875 Em:AV728994 Em:N91539
                    Em:BE906111 Em:BE372189 Em:AV725475 Em:AA229477
                    Em:BF031801 Em:AV659500 Em:AA357194 Em:AA075003
                    Em:AA228721 Em:M38714 Em:AA357193 Em:CO3105 Em:AA091429
                    Em:AM36129 Em:AM605144 Em:AM605142 Em:AM015656
                    Em:AA122673 Em:AI680140 Em:AA832355 Em:AA491183
                    Em:BE645240 Em:AA169452 Em:AM971488 Em:AM197669
                    Em:AA618518 Em:BE815910 Em:AA085746 Em:AI620439
                    Em:AI420298 Em:AI936423 Em:AA248996 Em:AA658199
                    Em:BE815919 Em:BE697434 Em:BE676861 Em:BE697435
                    Em:BE697436 Em:AM659002 Em:AM130169 Em:AI473966
                    Em:AM615082 Em:AM276216 Em:AM197858 Em:AM786005
                    Em:AA767728 Em:AA075034 Em:AA935887 Em:AA831190
```

```
Em:AI769381 Em:AI302136 Em:AM975096 Em:BE247082
Em:AI138958 Em:AA956621 Em:AM301242 Em:AM276204
Em:BE697440 Em:AA827775 Em:BE697442 Em:AI10858
Em:AA471313 Em:AI765568 Em:AV729224 Em:AM63377
Em:AA652231 Em:AM054987 Em:AI955586 Em:AA01305 Em:N92598
Em:AI445162 Em:BE048962 Em:AI955631 Em:AI15354
Em:AI955631 Em:BE754575 Em:AA79816 Em:AM630465
Em:AM197776 Em:AM467500"
/evidence=not experimental
polyA_site         complement(22369)
                    /gene="d4453H5.1"
polyA_site         complement(22371)
                    /gene="d4453H5.1"
polyA_site         complement(22372)
                    /gene="d4453H5.1"
polyA_site         complement(22373)
                    /gene="d4453H5.1"
polyA_site         complement(22375)
                    /gene="d4453H5.1"
polyA_signal       complement(22392..22397)
                    /gene="d4453H5.1"
CDS                complement(join(22850..23000,24209..24278,26185..26296,
                    26703..26795,28881..28978,30022..30129,31771..31925,
                    37238..37359,39165..39458,43686..43688))
                    /gene="d4453H5.1"
                    /codon_start=1
                    /evidence=not experimental
                    /product="d4453H5.1 (novel protein)"
                    /protein_id="CAC19504.1"
                    /db_xref="GI:11990034"
                    /db_xref="UniProt/TREMBL:Q9H1N7"
                    /translation="MDLTQDAKDIGNITVOETNNKNSBSIBCSKITMDLKNNSRKYI
                    SITVPSKQWTPHKSVDYVVLDMNLSKEKFLQPFICVAGVVPFLVGYLDEL
                    ISVBEKSGQWTLTIQVAPYSIPGLISLQIDKRRRIPGTTWITLFLYTMGSL
                    NTSGLYLPVPIYTRCKCLIPWMLGCFIQSKRNVDVSAIKMSGLIPLTADS
                    TTAEPNLTGVLLISLACADAVINQVBEKAKLNASNSKVLTSIGFYILGL
                    TCTSGIAPAVTECANPRTYGAFLFSLTGFGISFVLAIKIFGALIAVTTGK
                    AMTIVLSIFFAKPFPTFYVWSGLVLVIGIPLNVYSKNMDKIRLPSLYDLINKSEAR
                    KRTLAQTV"
repeat_region      24052..24185
                    /note="67 copies 2 mer aa 58% conserved"
repeat_region      24785..24986
                    /note="L2 repeat: matches 2477..2710 of consensus"
repeat_region      25084..25159
                    /note="MUT1A1 repeat: matches 107..186 of consensus"
repeat_region      25227..25288
                    /note="MUT1A1 repeat: matches 326..396 of consensus"
repeat_region      25296..25464
                    /note="MIR repeat: matches 2..188 of consensus"
repeat_region      27094..27187
                    /note="47 copies 2 mer aa 63% conserved"
repeat_region      27313..27538
                    /note="MIR repeat: matches 3..235 of consensus"
repeat_region      27770..27801
                    /note="16 copies 2 mer ac 93% conserved"
repeat_region      27772..27803
                    /note="8 copies 4 mer acac 93% conserved"
repeat_region      28055..28139
                    /note="8 copies 4 mer acac 93% conserved"
Query Match        13.1%; Score 262.4; DB 9; Length 92882;
                    Best Local Similarity 71.8%; Pred. No. 1.6e-51;
Matches 428; Conservative 0; Mismatches 156; Indels 12; Gaps 6;
QY                370 CTAAGAGTGTGGTTTACTAGAAATTA-ATGCTGAGCTGCTTATGCACTTCAGTTT 428
                    |||||
DB                50297 CCAGAGCACTGTCTGCATATAGAAATTAATTAATTAAGCCATGGAACCTATTAAATTTC 50356
                    |||||
QY                429 TAGTAGCCACATTAAACAGTAATAAAGGCTGGCGAGCTGCTACACCTGTATCCC 488
                    |||||
DB                50357 TGTCTAGTTACATTAAAGAGTAATAATGAGCGCGGTGTGCTGCTACAGCTCTATATCC 50416
                    |||||
QY                489 AACACTTGGAGGCTGAGGACGAGGACATCACCTTGTGTGACAGAGTTGAGACTAGCTG 548
                    |||||
```

Db 50417 AGCACTTTGGAGGCGCAAGGCGAATCACTTGAAGTCAAGATCCAGGCTG 50476
 QY 549 GCCAATCTGGCGAAACTCTGCTCTTA-AAAAAAAAATCAAAATTTAGCTGGCTGG 607
 Db 50477 GCCAATCTGGCGAAACTCTCTCTTAATAAAATTTAAACAGGCGCTGGTG 50536
 QY 608 CGGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACAAAGATCACTTAAACCAAG 667
 Db 50537 CGGGCGCTGTAATCTCAGCTGCTCAGAGGCGGAGACAAAGATCTTGAACCGC 50596
 QY 668 AGGTGAGGTTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 726
 Db 50597 AGGCAAGGTTCCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 50656
 QY 727 ACACCTTTGCTCAAAAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTAACT 786
 Db 50657 AGTCTGCTGCTCAAAAAAAGAGTGTG-ATATTAATTTTATTAATTAATTAAT 50712
 QY 787 TTAATAACCAATGTATCCCAATATCAATCAATCAATCAATCAATCAATCAATCAAT 846
 Db 50713 TATTAACTCAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 50771
 QY 847 ATGATGAGTACTTAACTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 906
 Db 50772 GTTAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 50827
 QY 907 ATATGTAATGCTGACAGCAATCTCAATTTGAGTACCTTCAAGTCTGAG 962
 Db 50828 GTATTTCACATTCACATTAATTTGGTTGGACAGCACATTTCAAGTCTGAG 50883

RESULT 15
 HS227L5 85304 bp DNA linear PRI 22-MAY-2002
 LOCUS Human DNA sequence from clone RP6-227L5 on chromosome Xp11.22-11.3,
 complete sequence.
 ACCESSION AL031585
 VERSION AL031585.2 GI:21212870
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 85304)
 REFERENCE
 AUTHORS Bird, C.
 TITLE Direct Submission
 JOURNAL Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquerry@sanger.ac.uk; clone requests: clonerequest@sanger.ac.uk
 COMMENT On May 25, 2002 this sequence version replaced gi:3980350.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL, Sw:
 SWISSPROT, Tr: TrEMBL, Wp: WormPep; Information on the WormPep
 database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
 was generated from part of bacterial clone contigs of human
 chromosome X, constructed by the Sanger Centre Chromosome X Mapping
 Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/chrX
 RP6-227L5 is from the library Rpci-6 constructed by the group of

Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pPAC4.
 FEATURES
 source
 Location/Qualifiers
 1. 85304
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="RZPD:RCP1P709L05227"
 /db_xref="taxon:9606"
 /chromosome="X"
 /map="p11.22-11.3"
 /clone="RP6-227L5"
 /clone_11b="Rpci-6"

ORIGIN
 Query Match 13.1%; Score 261.8; DB 9; Length 85304;
 Best Local Similarity 69.5%; Pred. No. 2.2e-51;
 Matches 448; Conservative 0; Mismatches 177; Indels 20; Gaps 6;
 373 GAGTACGCTGTTAGTAAATATGCTGAGCTGCTTATGCTATTTCCAGTTTATG 432
 Db 81598 GAAGATTTCTTTTGGAGCTTATTAATTTCTAGTATTAATTTAAATG 81657
 QY 433 AGCCACATTAACAGGTAA--AAAGCTGGCGGAGTGGCTCACACTGTAATCCAG 490
 Db 81658 TTAATTTTAAAAAGTAATTTCAAGCTGAGCATGCTGCTCACACTGTAATCTCAG 81717
 QY 491 CACTTGGAGGCTGAGGCGAGGAGATCACCTTTGCTGAGAGTTGAGACTGAGCTGGC 550
 Db 81718 CATTGGAGATGCTGAGGCGAGTGAATCACTTGAAGTCGGAGATTGAGACAGCTGGC 81777
 QY 551 CAACATGCGCAACTCTGCTCTTAAATAAATTAATTAATTAATTAATTAATTAATTAAT 610
 Db 81778 CAACATGCTGAACTCTATCTCT-ACATAAATTAATTAATTAATTAATTAATTAATTA 81836
 QY 611 GCGCTGTAATCTCAGCTGCTGAGAGGCGGAGACCAAGATCACTTAAACCAAGAGAG 670
 Db 81837 GTGCTGTATGCTCCAGTACTCAGAGGCTGAGGAGAGATTTGTTAAACCAAGAGG 81896
 QY 671 TGGAGTTGAGTGAAGCTGATGTCACCT-CACTCAACTGGAGAGCAAGTGACA 729
 Db 81897 CAGAGTTGAGTGAAGCAAGATCAACCACTGCTCAGCTTGAAGCAAGAGTGA 81956
 QY 730 CTTTGTCTCAAAAGAAAAAAGTAATTAATTAATTAATTAATTAATTAATTAATTA 783
 Db 81957 CTCAGTCTCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 82016
 QY 784 ACTTAAATCAACCAATGTATCCCAATTAATTAATTAATTAATTAATTAATTAATTA 843
 Db 82017 ACTTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 82074
 QY 844 ATTATGATGAGTACTTAACTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 903
 Db 82075 AATATGTTGAATTAATTAATTAATTT-----TTTCAATCAATGCTTTGAATGTCG 82126
 QY 904 TATATATGTAATGCTGACAGACATCTCAATTTGAGTACCTAATTTCAAGTCTCACT 963
 Db 82127 TGTGATTTTAACTTAACCAACATCTCAATTTGAATCAACCAATTAAGTCTCACT 82186
 QY 964 AGCCACATGCTGAGTGAAGTGAATGATGAGCAAGGATCTTA 1008
 Db 82187 AGCCACATGTAATCAAGGCTATCATTTTGAACAAAGATATGCTTA 82231

Search completed: February 26, 2005, 06:23:59
 Job time: 8539.79 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 13:36:01 ; Search time 1013.11 seconds
(without alignments)
11692.171 Million cell updates/sec

Title: US-09-936-271B-13_COPY_6000_8000

Perfect score: 2001
Sequence: 1 gacacccaagagcccccacg.....tcgcagggggaacttcta 2001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2001	100.0	11570	3	AAA95905
2	2001	100.0	11570	12	ADK52482
3	2001	100.0	11570	13	ADR72623
4	2001	100.0	11570	13	ADR72875
5	252.6	12.6	52216	6	AAH28355
6	252.6	12.6	52216	4	ABL50307
7	252.4	12.6	56098	11	ACN44922
8	240.2	12.0	55827	8	ACA60949
9	240.2	12.0	55827	10	ABX13671
10	240.2	12.0	58337	13	ADK56454
11	240.2	12.0	64423	13	ADK56462
12	235.6	11.8	27189	11	ACN45066
13	229.2	11.5	110000	11	ACN43998_3
14	227.6	11.4	2953	4	AAK68359
15	227.6	11.4	2953	4	AAK68360
16	227.6	11.4	3377	12	ADK64563
17	227.6	11.4	13744	4	AAK82406
18	227.6	11.4	13744	4	AAK68361
19	227.4	11.4	23071	4	AAK26699
20	227.4	11.4	23071	8	ADA98598

ALIGNMENTS

RESULT 1	AAA95905	standard; DNA; 11570 BP.
ID	AAA95905	
AC	AAA95905;	
DT	02-FEB-2001	(first entry)
DE	Human KLK-L2 gene.	
XX	Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;	
KW	Kallikrein-like protein; serine protease; cytosolic; cancer;	
XX	prostate cancer; ds.	
OS	Homo sapiens.	
XX		
PN	W0200053776-A2.	
PD	14-SEP-2000.	
PF	09-MAR-2000; 2000MO-CA000258.	
XX		
PR	11-MAR-1999; 99US-0124260P.	
PR	01-APR-1999; 99US-0127386P.	
PR	21-JUL-1999; 99US-0144919P.	
PA	(MOUN) MOUNT SINAI HOSPITAL.	
XX		
PI	Yousef GM, Diamandis EP;	
XX	WPI; 2000-587440/55.	
DR	P-PSDB; AAB21296.	
XX		
PT	New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L	
XX	protein mediated disorders, especially cancer.	
PS	Claim 1; Page 143-149; 184pp; English.	
XX		
CC	The present sequence is the coding sequence of the human KLK-L2 gene,	
CC	which encodes a kallikrein-like protein. Kallikreins and kallikrein-like	
CC	proteins are a subgroup of the serine protease enzyme family. They	
CC	catalyse the selective cleavage of specific polypeptide precursors to	
CC	release peptides with potent biological activity. Nucleic acids encoding	

C	21	227.4	11.4	23071	8	ABX74048	ABX74048 Human nov
	22	227.4	11.4	23071	8	ADA44323	Ada44323 Human sec
	23	227.4	11.4	23071	10	ADK20712	Adk20712 Human sec
	24	227.4	11.4	23071	10	ADK10881	Adk10881 Human sec
	25	226.8	11.3	12275	4	AAK85734	Aak85734 Human imm
	26	226.2	11.3	133955	11	ACN45170	Acn45170 Human gen
	27	225.4	11.3	325791	4	AAK43104	Aak43104 Human Oes
	28	225.2	11.3	35641	6	ABL64428	Ab164428 Stomach c
	29	225.2	11.3	35641	6	ABN95727	Abn95727 Gene #222
	30	224.8	11.2	629	12	ADM99987	Adm99987 Human STR
	31	224.8	11.2	629	12	ADM99988	Adm99988 Human STR
	32	223.8	11.2	15765	5	ABA18276	Abal18276 Human ner
	33	223.8	11.2	15765	5	ABA19009	Abal19009 Human ner
	34	223.8	11.2	15765	5	ABA18250	Abal18250 Human ner
	35	223.8	11.2	15765	5	ABA18250	Abal18250 Human ner
	36	223.6	11.2	12657	4	AAK85733	Aak85733 Human imm
	37	223.2	11.2	110000	10	ADG70447_3	Adg70447_3
	38	223.2	11.2	110000	10	ABK79565_3	Abk79565_3
	39	222	11.1	72409	13	ABD33061	Abd33061 Human can
	40	221.6	11.1	2591	4	AAH16224	Aah16224 Human can
	41	221.6	11.1	2591	12	ADK22885	Adk22885 Human MP2
	42	221.6	11.1	2591	13	ADR14082	Adr14082 Human NF-
	43	220.8	11.0	66685	4	AAK07380	Aak07380 Human gen
	44	220.8	11.0	66685	6	ABK73149	Abk73149 Human CLA
	45	220.8	11.0	96587	9	ADA02984	Ada02984 Human MAP

CC kallikrein-like proteins KLK-11, KLK-12, KLK-13, KLK-14, KLK-15 and KLK-16 have been isolated. The proteins are useful in the treatment, monitoring and diagnosis of cancers, especially prostate cancer. They can also be used to identify a substance that can associate with or mediate the biological activity of the proteins. Antibodies can be used to treat conditions mediated by the kallikrein-like proteins

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 U; 0 Other;

Query Match 100.0%; Score 2001; DB 3; Length 11570;

Best Local Similarity 100.0%; Pred. No.0;

Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GACAAACAAGAGCCCCCAAGAGTGTCCAGGTTCTTCTTGAATACCAACCATCTCTGC 60
DB 6000 GACAAACAAGAGCCCCCAAGAGTGTCCAGGTTCTTCTTGAATACCAACCATCTCTGC 6055
QY 61 CGGCTTCATCTTCTTCACCTTCTCATGTGTCTCTGTTTGAAGTGCACCTTCCCTAAG 120
DB 6060 CGGCTTCATCTTCTTCACCTTCTCATGTGTCTCTGTTTGAAGTGCACCTTCCCTAAG 6119
QY 121 TCCTCAGGCTTGAATATCAGGCTGTAGTCAAGAAAGGTCGAGAGTCTTACCCA 180
DB 6120 TCCTCAGGCTTGAATATCAGGCTGTAGTCAAGAAAGGTCGAGAGTCTTACCCA 6179
QY 181 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTGCGCAG 240
DB 6180 GACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCTGCGCAG 6239
QY 241 TGAAGACACCTCTCTTATTACAGAGATACACTGATGTCCTGCTGTTACATGAGAC 300
DB 6240 TGAAGACACCTCTCTTATTACAGAGATACACTGATGTCCTGCTGTTACATGAGAC 6299
QY 301 GTTGCAAAATTCGAGATTCAGCAATTCGCAAGATGCAAGACCCCTGTTCTCACA 360
DB 6300 GTTGCAAAATTCGAGATTCGCAAGATGCAAGACCCCTGTTCTCACA 6359
QY 361 GCTCATACCTTAGAGTGTGTGTATGAAATATGCTGAGTCTTATGCTATTC 420
DB 6360 GCTCATACCTTAGAGTGTGTGTATGAAATATGCTGAGTCTTATGCTATTC 6419
QY 421 CAGTTTTTATGATGCCAATTAAACAGGTAAGGCTGGCGCAATGCTCACACT 480
DB 6420 CAGTTTTTATGATGCCAATTAAACAGGTAAGGCTGGCGCAATGCTCACACT 6479
QY 481 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGAGATCACTTTGTCAGAGTTTGAA 540
DB 6480 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGAGATCACTTTGTCAGAGTTTGAA 6539
QY 541 CTAGCTGGCAACATGCGGAAACTGTCTCTTAAATAAATAATTAAGCTGGC 600
DB 6540 CTAGCTGGCAACATGCGGAAACTGTCTCTTAAATAAATAATTAAGCTGGC 6599
QY 601 ATGCTGGCGGCGCTGTATCTCAGCTGCTCAGAGCGGAGACACAAAGATCACTTAA 660
DB 6600 ATGCTGGCGGCGCTGTATCTCAGCTGCTCAGAGCGGAGACACAAAGATCACTTAA 6659
QY 661 ACCCAGAGAGTGGAGGTGCAATGAGCTGAGATCGGCACTGACCTCCAACTGGGAGAC 720
DB 6660 ACCCAGAGAGTGGAGGTGCAATGAGCTGAGATCGGCACTGACCTCCAACTGGGAGAC 7219
QY 721 AGAGTGACACTTGTCTCAAAAAGAAAAAACAATGTAAGAAAGAAACAGTGAAG 780
DB 6720 AGAGTGACACTTGTCTCAAAAAGAAAAAACAATGTAAGAAAGAAACAGTGAAG 6779
QY 781 TTAACTTTAATAACCAATGATCCCAATACATCAATTTCAAGTGTATTAATAA 840
DB 6780 TTAACTTTAATAACCAATGATCCCAATACATCAATTTCAAGTGTATTAATAA 6839
QY 841 ACAATTAATAGATGATCTTATCATTTCTTTCTTTTCAATTAAGTCTTGAAGT 900
DB 6840 ACAATTAATAGATGATCTTATCATTTCTTTCTTTTCAATTAAGTCTTGAAGT 6899
```

```
QY 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATGACTAATTTGAGTGTCTC 960
DB 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGACATGACTAATTTGAGTGTCTC 6959
QY 961 AGTAGCCACATGTGGCTAGCAAGTACTGTATTTGATGCGGAGATTTAGAGGAAAGATC 1020
DB 6960 AGTAGCCACATGTGGCTAGCAAGTACTGTATTTGATGCGGAGATTTAGAGGAAAGATC 7019
QY 1021 AAGGCTGTTTTATGTTGTTGGCAGGTTGTGACATGCAATTAAGATTCATTAATAG 1080
DB 7020 AAGGCTGTTTTATGTTGTTGGCAGGTTGTGACATGCAATTAAGATTCATTAATAG 7079
QY 1081 GGCACCTCCGTGTACAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGTAGTTAACTGTC 1140
DB 7080 GGCACCTCCGTGTACAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGTAGTTAACTGTC 7139
QY 1141 TTGTTTCAACAAATCTGTATATATGACAGTCTTTCTAGCAAGTGTGTATTTCTTGA 1200
DB 7140 TTGTTTCAACAAATCTGTATATATGACAGTCTTTCTAGCAAGTGTGTATTTCTTGA 7199
QY 1201 GGAGAGAAAGAGAAATCTGTAGGTAATTTTACAAGAGATATTAATACAGGGGATTA 1260
DB 7200 GGAGAGAAAGAGAAATCTGTAGGTAATTTTACAAGAGATATTAATACAGGGGATTA 7259
QY 1261 ATTGCAAGCTGCTGGAAGGCTGGAAGAACAGTTAAATAAATAAATACTCTGTGTC 1320
DB 7260 ATTGCAAGCTGCTGGAAGGCTGGAAGAACAGTTAAATAAATAAATACTCTGTGTC 7319
QY 1321 AAGATCTGATTAATAGGCAATTTCAAGAGTGTATTAAGTTAACTCCCAATTAATAC 1380
DB 7320 AAGATCTGATTAATAGGCAATTTCAAGAGTGTATTAAGTTAACTCCCAATTAATAC 7379
QY 1381 ATGTTTTATGATAGTAAACAAATTAAGGCGCAATATTTCAAAAAGGTTGAGGGAGCTC 1440
DB 7380 ATGTTTTATGATAGTAAACAAATTAAGGCGCAATATTTCAAAAAGGTTGAGGGAGCTC 7439
QY 1441 CTTGAGAGGTGGCAATTTGAGAGAGAAATGATGACACAAAGATCTGTAAGT 1500
DB 7440 CTTGAGAGGTGGCAATTTGAGAGAGAAATGATGACACAAAGATCTGTAAGT 7499
QY 1501 TTTAAGGGGAAAGAAAGGCACTGCAAGAGGCTTGAGGCAATTAAGGATTTGCTGATTC 1560
DB 7500 TTTAAGGGGAAAGAAAGGCACTGCAAGAGGCTTGAGGCAATTAAGGATTTGCTGATTC 7559
QY 1561 AAAAGAAAGAGAAACCAATGCACTGAGAGAAACAAAGTGGGGGCAACATGAGAAAGT 1620
DB 7560 AAAAGAAAGAGAAACCAATGCACTGAGAGAAACAAAGTGGGGGCAACATGAGAAAGT 7619
QY 1621 ACGCTGAGAGTGTAGGCAAGGGCGAATGCTCTGCAAGTATTTCTTGTCACCAACAGA 1680
DB 7620 ACGCTGAGAGTGTAGGCAAGGGCGAATGCTCTGCAAGTATTTCTTGTCACCAACAGA 7679
QY 1681 GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTGAGAAAGCAGAAATTTAAATCAAC 1740
DB 7680 GCTTCCCTATGTTCTTAATGGAAGCTGTATCTGTGAGAAAGCAGAAATTTAAATCAAC 7739
QY 1741 TGTTCATCAACACAGACCCCTCTCTGTATGAGGCTCCCAAGGGATCTAGAAAGAGCTA 1800
DB 7740 TGTTCATCAACACAGACCCCTCTCTGTATGAGGCTCCCAAGGGATCTAGAAAGAGCTA 7799
QY 1801 AGTTAACAGACTCTCATTAAGCAGGAGTGTGTTTCAACAGTATTAAGAGCTGGGAGAT 1860
DB 7800 AGTTAACAGACTCTCATTAAGCAGGAGTGTGTTTCAACAGTATTAAGAGCTGGGAGAT 7859
QY 1861 CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCCCAATTTGTAACAACTGACACA 1920
DB 7860 CAGAGTACTCCAGTCCCATGCTATGAAAGCTCCCCCAATTTGTAACAACTGACACA 7919
QY 1921 ATGCAACACCTCCAGCTCTCCCATTTCTCTGAGGCTCTGGGAGTGGGAGTGGG 1980
DB 7920 ATGCAACACCTCCAGCTCTCCCATTTCTCTGAGGCTCTGGGAGTGGGAGTGGG 7979
QY 1981 TTGCGAGGGGGAACCTTTTA 2001
```


Db 7980 TTGCGAGGGGAAAACCTTTA 8000

RESULT 2
ADK52482
ID ADK52482 standard; DNA; 11570 BP.
XX
AC ADK52482;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human kallikrein 5 encoding sequence.
XX
KM kallikrein 5; cancer; Cytostatic; Immunostimulant; db.
XX
OS Homo sapiens.
XX
PN MO2004021008-A2.
XX
PD 11-MAR-2004.
XX
PF 28-AUG-2003; 2003MO-CA001310.
XX
PR 28-AUG-2002; 2002US-0407333P.
XX
PA (MOUN) MOUNT SINAI HOSPITAL.
XX
PI (YOUS/) YOUSEF G.
XX
PI Diamandis BP;
XX
DR MPI; 2004-239232/22.
XX
P-PSDB; ADK52481.
XX
PT Detecting kallikrein 5 associated with cancer, useful in diagnosing,
PT monitoring, detecting, imaging and treating breast or ovarian carcinoma,
PT compares comparing the detected amount of kallikrein 5 in a sample with
PT a standard sample.
XX
PS Disclosure; SEQ ID NO 2; 59pp; English.
XX
XX The present invention relates to detecting kallikrein 5 associated with
CC breast or ovarian cancer in a patient comprises detecting in the sample
CC kallikrein 5 and comparing the detected amount with an amount detected
CC for a standard. The method is useful in detecting kallikrein 5 associated
CC with breast or ovarian cancer in a patient. The methods and kits are
CC useful in diagnosing, monitoring, detecting, imaging and treating breast
CC or ovarian carcinoma. The kallikrein is useful in preparing a vaccine for
CC preventing and treating breast and ovarian cancer and for stimulating or
CC enhancing antibody production or for inducing an immune response. The
CC present sequence represents human kallikrein 5 encoding sequence.
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 2001; DB 12; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACAAACAAGAGCCCCCAAGTAGTGTCCAGGTTCTTCTTGAATACCGACCACTCTCTGC 60
Db 6000 GACAAACAAGAGCCCCCAAGTAGTGTCCAGGTTCTTCTTGAATACCGACCACTCTCTGC 6059

QY 61 CGGCTTCATCTTTCTTCATCTTCTCATTTGTTCTGTTTGAAGTCACTTCCCTAAG 120
Db 6060 CGGCTTCATCTTTCTTCATCTTCTCATTTGTTCTGTTTGAAGTCACTTCCCTAAG 6119

QY 121 TCTCCAGTCTGAATATCAGGTGCTAATGAGAAAAGTGAGAGATCTTACCCGA 180
Db 6120 TCTCCAGTCTGAATATCAGGTGCTAATGAGAAAAGTGAGAGATCTTACCCGA 6179

QY 181 GACAGATAGATGACACCATCTTCTGCCTGAGTGAAGAGAGATAGAGATCTCTGCAGG 240
Db 6180 GACAGATAGATGACACCATCTTCTGCCTGAGTGAAGAGAGATAGAGATCTCTGCAGG 6239

QY 241 TGAGACACCTCTCTTATTTACAGCAGATACACACTGAGTGCCAACTCGGTAACTGAGG 300
Db 6240 TGAGACACCTCTCTTATTTACAGCAGATACACACTGAGTGCCAACTCGGTAACTGAGG 6299

QY 301 GTTGCCAAATTTCTGAGAAATCCAGCAATTCGCAAGACGTCAAGACCCTGTTCTCAGA 360
Db 6300 GTTGCCAAATTTCTGAGAAATCCAGCAATTCGCAAGACGTCAAGACCCTGTTCTCAGA 6359

QY 361 GCTCATACCTCTGAGATAGTGTGTTTATGAGAAATATGCTGAGCTGTTATGCTATTC 420
Db 6360 GCTCATACCTCTGAGATAGTGTGTTTATGAGAAATATGCTGAGCTGTTATGCTATTC 6419

QY 421 CAGTTTTTATGATGCCAATTTAAACAGGTAAATAAGGCTGGCGGAGTGCTCACACCT 480
Db 6420 CAGTTTTTATGATGCCAATTTAAACAGGTAAATAAGGCTGGCGGAGTGCTCACACCT 6479

QY 481 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAAGAGTTTGAGA 540
Db 6480 GTAATCCAGCACTTTGGAGGCTGAGGAGGAGGAGATCACTTTGGTCAAGAGTTTGAGA 6539

QY 541 CTAGCTGGCCAACTATGGGAAACTCTGTCTCTTAAAAAAATATCAAAAATTAGCCTGGC 600
Db 6540 CTAGCTGGCCAACTATGGGAAACTCTGTCTCTTAAAAAAATATCAAAAATTAGCCTGGC 6599

QY 601 ATGCTGGCGGCGGCGGTATATCTCAGCTGCTCAGAGGCGGAGACCAAGAATCACTTAA 660
Db 6600 ATGCTGGCGGCGGCGGTATATCTCAGCTGCTCAGAGGCGGAGACCAAGAATCACTTAA 6659

QY 661 ACCCAGAGGTGAGGTGCTCAGTGAAGTGAAGTGTGCTGCTCACTCACTCAACTGGAGAC 720
Db 6660 ACCCAGAGGTGAGGTGCTCAGTGAAGTGAAGTGTGCTGCTCACTCACTCAACTGGAGAC 6719

QY 721 AGAGTGAACCTTTGTCTCAAAAAAGAAAAAAGAAAAAAGAAAAAGGTTGAG 780
Db 6720 AGAGTGAACCTTTGTCTCAAAAAAGAAAAAAGAAAAAAGAAAAAGGTTGAG 6779

QY 781 TTAACCTTAAATACCCAAATGTAATCCCAATATCAATATTTCAAGGTATTTAATATAA 840
Db 6780 TTAACCTTAAATACCCAAATGTAATCCCAATATCAATATTTCAAGGTATTTAATATAA 6839

QY 841 ACAATATGATGATGATTAATCTTACATCTTTCTTCTTTCATATTAAGTCTTGAAGT 900
Db 6840 ACAATATGATGATGATTAATCTTACATCTTTCTTCTTTCATATTAAGTCTTGAAGT 6899

QY 901 GAGTATATATGTTATGCTGACAGACATCTCAATTTGAGACTGATCACTTCAAGTGTCT 960
Db 6900 GAGTATATATGTTATGCTGACAGACATCTCAATTTGAGACTGATCACTTCAAGTGTCT 6959

QY 961 AGTAGCCATATGCTGCTAGCAGTACTGTAATGGAATGGACCGAATCTAAGAGAAAGATC 1020
Db 6960 AGTAGCCATATGCTGCTAGCAGTACTGTAATGGAATGGACCGAATCTAAGAGAAAGATC 7019

QY 1021 AGGCTGTTTTTGTATGATGTTGGGAGGTTGTCAGTCACTATTAAGATATTAATAG 1080
Db 7020 AGGCTGTTTTTGTATGATGTTGGGAGGTTGTCAGTCACTATTAAGATATTAATAG 7079

QY 1081 GGCACCTCGGTATACAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGATGTTAAGTGC 1140
Db 7080 GGCACCTCGGTATACAGATGTCAGTTTGGCAGTTTTCAGGCGTGTGATGTTAAGTGC 7139

QY 1141 TTGTTTCAACAAAATCTGTAATATGACATTTTCTGCAAGTGTGTAATAATCTCTGA 1200
Db 7140 TTGTTTCAACAAAATCTGTAATATGACATTTTCTGCAAGTGTGTAATAATCTCTGA 7199

QY 1201 GGAAGAAAAGAGAAATCTGTATGATTTTATCAAGAGAAATTTAATACAGGGATTA 1260
Db 7200 GGAAGAAAAGAGAAATCTGTATGATTTTATCAAGAGAAATTTAATACAGGGATTA 7259

QY 1261 ATTGCAAGCTGTGGAAGGGCTGAGAGAACAAAGTTAAAAAATAAAAAATCTGTGTC 1320
Db 7260 ATTGCAAGCTGTGGAAGGGCTGAGAGAACAAAGTTAAAAAATAAAAAATCTGTGTC 7319

OY		1321	AAGATCTGCATTAATTAGGGCAATTTGACAGGTGTGAAGGTTAAACCCAAAATPAAAC	1380
Db		7320	AAGAATCTGCATAAATTAGGCCAATTTTTCAGAGTGCTTAAAGTTAACCCAAAATPAAAC	7379
OY		1391	ATGCTTTTAGATAGTAAACAATAAGGGCCAATATTTCAAAAAAGTGCTCAGGGAGCCTC	1440
Db		7380	ATGTTTTTAGATAGTAAACAATAAGGCCAATATTTCAAAAAAGTGCTCAGGGAGCCTC	7439
OY		1441	CTTGGAGAGGTGGCATTTTGAGCAGAGATGAGTGAACAAMAGAGCTAAACTCGTAAGT	1500
Db		7440	CTTGGAGAGGTGGCATTTTGAGCAGAGATGAGTGAACAAMAGAGCTAAACTCGTAAGT	7499
OY		1501	TTAAGGGGAAAGAAAGGACGCTGCMAAGGCCCTGAGGACAGTAAGGAATTTGGCTGATTC	1560
Db		7500	TTAAGGGGAAAGAAAGGACGCTGCMAAGGCCCTGAGGACAGTAAGGAATTTGGCTGATTC	7559
OY		1561	AAAGAAAGAAAGAAACCATATGCACTGGAAGAACAAAGTGCGGCGCACAGTAGAAAGTG	1620
Db		7560	AAAGAAAGAAAGAAACCATATGCACTGGAAGAACAAAGTGCGGCGCACAGTAGAAAGTG	7619
OY		1621	AACCTGGAGGTGTAGGCGAGGGCCGAATGCTCTGCAAGTATTTCTTGTCACCAACACAGA	1680
Db		7620	AACCTGGAGGTGTAGGCGAGGGCCGAATGCTCTGCAAGTATTTCTTGTCACCAACACAGA	7679
OY		1681	GCTTCCCATGTGTTCTAATGGAAGCTGATCTGTTGAGGAAGACGAATTTAATTAACAAC	1740
Db		7680	GCTTCCCATGTGTTCTAATGGAAGCTGATCTGTTGAGGAAGACGAATTTAATTAACAAC	7739
OY		1741	TGTTACATCAACACAGACCCCTTCTCTGTATTCAGGCTCCCAAGGATCTAGAAAGACGTA	1800
Db		7740	TGTTACATCAACACAGACCCCTTCTCTGTATTCAGGCTCCCAAGGATCTAGAAAGACGTA	7799
OY		1801	AGTTAAACAAGCTCTCATTAGCAGGGGTGTGTGTTCAACAGTAGTTAGAGCTGGGATTT	1860
Db		7800	AGTTAAACAAGCTCTCATTAGCAGGGGTGTGTGTTCAACAGTAGTTAGAGAGCTGGGATTT	7859
OY		1861	CAGAGTAGTCTCCAGTCCCATGGCTATGAAAGCTCCCCCAAATTTGTAACAACCTGACAA	1920
Db		7860	CAGAGTAGTCTCCAGTCCCATGGCTATGAAAGCTCCCCCAAATTTGTAACAACCTGACAA	7919
OY		1921	ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCCCTGTGGGGGGGTGGG	1980
Db		7920	ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTGTGTGCCCTGTGGGGGGGTGGG	7979
OY		1981	TTGCTAGGGGGGAAAACCTTTTA 2001	
Db		7980	TTGCGAGGGGGGAAAACCTTTTA 8000	
RESULT 3				
ADR72623				
ID	ADR72623	standard; DNA; 11570 BP.		
XX	AC			
XX	ADR72623;			
DT	02-DEC-2004	(first entry)		
DE	Human renal cell carcinoma-related kallikrein 5 (KLK5) DNA 2.			
KM	kallikrein 5; renal cell carcinoma; cytostatic; tumour growth;			
KW	cancer metastasis; chemotherapy; human; serine protease;			
XX	chromosome 19q13.4; KLK5; ds; gene.			
OS	Homo sapiens.			
FH	Key	Location/Qualifiers		
FT	CDS	2221..11247		
FT		/*tag= b		
FT	/product=	"Human renal cell carcinoma-related kallikrein		
FT		5 (HK5) protein"		
FT	exon	2221..2293		
FT		/*tag= a		
FT		/number= 1		

Query Match	100.0%;	Score 2001;	DB 13;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Query Match	100.0%;	Score 2001;	DB 13;	Length 11570;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

XX kallikrein 5; tumour marker; ovarian cancer;
KM epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
XX KLK5; ds; gene.
OS Homo sapiens.
XX
XX
FH Location/Qualifiers
FT CDS 2221..11247
FT /tag= b
FT /product= "Human ovarian cancer-related tumour marker
FT kallikrein 5 (hK5) protein"
FT exon /tag= a
FT /number= 1
FT 2294..4761
FT /tag= c
FT /number= 1
FT 4762..5023
FT /tag= d
FT /number= 2
FT 5024..5762
FT /tag= e
FT /number= 2
FT 5763..6019
FT /tag= f
FT /number= 3
FT 6020..6104
FT /tag= g
FT /number= 3
FT 6105..6238
FT /tag= h
FT /number= 4
FT 6239..11091
FT /tag= i
FT /number= 4
FT 11092..11247
FT /tag= j
FT /number= 5
PN MO2004075713-A2.
XX
XX 10-SEP-2004.
PD
XX
PF 26-FEB-2004; 2004MO-CA000281.
XX
PR 26-FEB-2003; 2003US-0450406P.
XX
PA (MOUN) MOUNT SINAI HOSPITAL.
XX
XX Diamandis EP;
PI
XX MPI: 2004-661815/64.
DR P-PSDB; ADR12873.
XX
XX kallikrein markers detection method for detecting ovarian cancer in
PT patient, involves detecting kallikrein markers and CA125 in sample
PT obtained from patient, and comparing detected amounts with standard
PT amounts.
XX
XX Example 2; SEQ ID NO 5; 102pp; English.
XX
XX The invention relates to a novel method for detecting a plurality of
CC kallikrein markers associated with ovarian cancer. The method comprises
CC obtaining a sample from a patient and detecting in the sample a plurality
CC of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
CC the kallikrein markers are selected from the group consisting of
CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
CC kallikrein 11. The detected amounts of the kallikrein markers are
CC compared with standard amounts. The method of the invention may be useful
CC for detecting kallikrein markers associated with ovarian cancer in a
CC patient and thus for detecting ovarian cancer, particularly epithelial
CC ovarian carcinoma. The current sequence is that of the human ovarian

CC cancer-related tumour marker kallikrein 5 (KLK5) DNA 2 of the invention
CC which encodes a secreted serine protease and is located at chromosome
CC 19q13.4.
XX
XX
SQ Sequence 11570 BP; 3034 A; 2562 C; 3325 G; 2649 T; 0 U; 0 Other;
Query Match 100.0%; Score 2001; DB 13; Length 11570;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACAAACCAAGAGCCCCCAAGAGTGCAGAGTCTTCTTATACCAATCCATCTGCG 60
DB 6000 GACAAACCAAGAGCCCCCAAGAGTGCAGAGTCTTCTTATACCAATCCATCTGCG 6059
QY 61 CGCCTTCATCTTCTTCACCTCTCATTTGTTCTGTTTGAAGTGCATTCCTTAAGG 120
DB 6060 CGCCTTCATCTTCTTCACCTCTCATTTGTTCTGTTTGAAGTGCATTCCTTAAGG 6119
QY 121 TCCTCCAGTGTGAATATCAGCGTCTAAGTCAAGAAAAGTGGAGAGATGCTTACCCGA 180
DB 6120 TCCTCCAGTGTGAATATCAGCGTCTAAGTCAAGAAAAGTGGAGAGATGCTTACCCGA 6179
QY 181 GACAGATAGTATGACACCATGTTCTGCGCGGTGACAAAGCAGTGAAGCTCTGCGCAGG 240
DB 6180 GACAGATAGTATGACACCATGTTCTGCGCGGTGACAAAGCAGTGAAGCTCTGCGCAGG 6239
QY 241 TGAGGACACCTCTCTTATTCAGCAGATACACACTGAGTCCATCTGGTAAATGAGAC 300
DB 6240 TGAGGACACCTCTCTTATTCAGCAGATACACACTGAGTCCATCTGGTAAATGAGAC 6299
QY 301 GTTGCCAAATTCGTGAATATCCAGCAATTGCCAAGACAGTCAAGACCCCTGTTCTACAGA 360
DB 6300 GTTGCCAAATTCGTGAATATCCAGCAATTGCCAAGACAGTCAAGACCCCTGTTCTACAGA 6359
QY 361 GCTCATACCTTGAATAGTGTGTTTATGAAATAATGCTGATATGATATTC 420
DB 6360 GCTCATACCTTGAATAGTGTGTTTATGAAATAATGCTGATATGATATTC 6419
QY 421 CAGTTTTTATGAGCCACTTAATAACAGGTAAATAAGCTGGCCAGTGGCTCACACT 480
DB 6420 CAGTTTTTATGAGCCACTTAATAACAGGTAAATAAGCTGGCCAGTGGCTCACACT 6479
QY 481 GTAATCCAGACCTTTGGAGGCTGAGGCGAGGACATCACTTTGTCAGAGTTTGAGA 540
DB 6480 GTAATCCAGACCTTTGGAGGCTGAGGCGAGGACATCACTTTGTCAGAGTTTGAGA 6539
QY 541 CTAGCCTGGCCAAACATGGGAAACTCTGCTCTAATAAAAAAATATCAAAATATAGCTGGC 600
DB 6540 CTAGCCTGGCCAAACATGGGAAACTCTGCTCTAATAAAAAAATATCAAAATATAGCTGGC 6599
QY 601 ATGATGAGCGAGGCGCTGTATCTCAGCTGCTCAGAGGCGCGAGACACAAATCACTTAA 660
DB 6600 ATGATGAGCGAGGCGCTGTATCTCAGCTGCTCAGAGGCGCGAGACACAAATCACTTAA 6659
QY 661 ACCCAGAGAGTGGAGGTTGACGTGAGTGCATGATCGTCCATCTCAACTGGAGAC 720
DB 6660 ACCCAGAGAGTGGAGGTTGACGTGAGTGCATGATCGTCCATCTCAACTGGAGAC 6719
QY 721 AAGATGACACTTTTGTCCAAAAGAAAAAAGAAAAAAGAAAAAGAAAGAGGGAAG 780
DB 6720 AAGATGACACTTTTGTCCAAAAGAAAAAAGAAAAAAGAAAAAGAAAGAGGGAAG 6779
QY 781 TTAACCTTATATACCAATGTATCCCAATACATATTCATTAATATATATAA 840
DB 6780 TTAACCTTATATATACCAATGTATCCCAATACATATTCATTAATATATATAA 6839
QY 841 ACAAT 900
DB 6840 ACAAT 6899
QY 901 GAGTAT 960
DB 6900 GAGTAT 6959

QY	961	AGTAAAGCCACATGTGGCTGACAGTTACCTGTAATTGACATGGCCAGGATCTTACAGGAAAAATC	1020
Db	6960	AGTAAAGCCACATGTGGCTGACAGTTACCTGTAATTGACATGGCCAGGATCTTACAGGAAAAATC	7019
QY	1021	AGGCGTGTATTTGTATAGTGTGGCAGGTGTGCATGCACTAAAGATACCATATCTTAATAG	1080
Db	7020	AGGCGTGTATTTGTATAGTGTGGCAGGTGTGCATGCACTAAAGATACCATATCTTAATAG	7079
QY	1081	GGCACTCCGTGTTACAGATGTCAGTTTGGCAGTTTCAAGCGTGTGTGTAAGTGC	1140
Db	7080	GGCACTCCGTGTTACAGATGTCAGTTTGGCAGTTTCAAGCGTGTGTGTAAGTGC	7139
QY	1141	TTGTTTCAACAAAATCTGTAATATATACAGTTTTTCAGCAAGGCGTGGTAAATATCTTGA	1200
Db	7140	TTGTTTCAACAAAATCTGTAATATATACAGTTTTTCAGCAAGTGTGTGTAATATCTTGA	7199
QY	1201	GGAAAGAAAAGAAAATCTGTGGTATTTTTCAGAGAAATATTTAATACAGGGATTA	1260
Db	7200	GGAAAGAAAAGAAAATCTGTGGTATTTTTCAGAGAAATATTTAATACAGGGATTA	7259
QY	1261	ATTGCAAAAGCTGTGGAAAGGCTGTGAGGAACAAACTTAAAAATTAATAACTGTGCTC	1320
Db	7260	ATTGCAAAAGCTGTGGAAAGGCTGTGAGGAACAAACTTAAAAATTAATAACTGTGCTC	7319
QY	1321	AAGAATCTGCATTAATATAGGCGCAATTTCAAGAGTGTAAAGTTAAACCCCAAAATTAAC	1380
Db	7320	AAGAATCTGCATTAATATAGGCGCAATTTCAAGAGTGTAAAGTTAAACCCCAAAATTAAC	7379
QY	1381	ATGTTTTAGGATAGTAAACAATAAGGCCAATAATTTCAAAAAGTGTGTCAGGGAGCCTC	1440
Db	7380	ATGTTTTAGGATAGTAAACAATAAGGCCAATAATTTCAAAAAGTGTGTCAGGGAGCCTC	7439
QY	1441	CTTGAGAGAGTGTGGCAATTTGAGCAGAGATGAGTGCACAAAGAAGCTAAATCTGTGAAGT	1500
Db	7440	CTTGAGAGAGTGTGGCAATTTGAGCAGAGATGAGTGCACAAAGAAGCTAAATCTGTGAAGT	7499
QY	1501	TTAAAGGGAAAAGAAAAGGCAAGTGCMAAGGCCCTGTAGGCAAGTAAAGAAATTTGGCTGATTC	1560
Db	7500	TTAAAGGGAAAAGAAAAGGCAAGTGCMAAGGCCCTGTAGGCAAGTAAAGAAATTTGGCTGATTC	7559
QY	1561	AAAAGAAAGAGGAAACCAATGCAACATGAGAACAAAGATGGGGGCAACATAGAAAGTG	1620
Db	7560	AAAAGAAAGAGGAAACCAATGCAACATGAGAACAAAGATGGGGGCAACATAGAAAGTG	7619
QY	1621	ACGCTGAGAGTGTATAGSCAGGGGCGAATCTCTGCAGATATTTCTTGGTCAACAACAGA	1680
Db	7620	ACGCTGAGAGTGTATAGSCAGGGGCGAATCTCTCTGCAGATATTTCTTGGTCAACAACAGA	7679
QY	1681	GCTTCCCTATATGTTCTAATGAAAGCTGTATCTGTTGAGAGAAACAGAAATTTAAATCAAC	1740
Db	7680	GCTTCCCTATATGTTCTAATGAAAGCTGTATCTGTTGAGAGAAACAGAAATTTAAATCAAC	7739
QY	1741	TGTTATCATCAACAGCAACCTCTCTGTATTCAGGCTCCCAAGGGATTTAGAAAGACGTA	1800
Db	7740	TGTTATCATCAACAGCAACCTCTCTGTATTCAGGCTCCCAAGGGATTTAGAAAGACGTA	7799
QY	1801	AGTTAAACAAGCTTCTATTAGCAGGGTGTGTGTTTCAACAGTAAAGTTAGAGACTGGGATTT	1860
Db	7800	AGTTAAACAAGCTTCTATTAGCAGGGTGTGTGTTTCAACAGTAAAGTTAGAGACTGGGATTT	7859
QY	1861	CAGAGATCTCAAGTCCCAATGGCTATGAAAAGCTCCCCCAAAATTTGTAACAACCTGCACA	1920
Db	7860	CAGAGATCTCAAGTCCCAATGGCTATGAAAAGCTCCCCCAAAATTTGTAACAACCTGCACA	7919
QY	1921	ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTGTGCCCTGGGTGTGGGGGGGTGGG	1980
Db	7920	ATGCAACAACCTCCCAAGCTCTCCCAATTTCTTCTGTGCCCTGGGTGTGGGGGGGTGGG	7979
QY	1981	TTGCGAAGGGGAAAACTTTTA 2001	
Db	7980	TTGCGAAGGGGAAAACTTTTA 8000	

DB	AAH28355	standard; DNA, 52216 BP.
ID	AAH28355	
XX	AAH28355;	
AC	AAH28355;	
XX		
DT	05-SEP-2001	(first entry)
XX		
DE	Nucleotide sequence of the human muasashi promoter.	
XX		
KW	Muasashi promoter; multipotential neural progenitor cell;	
XX	neural stem cell; central nervous system; ss.	
OS	Homo sapiens.	
PN	WO200146384-A2.	
PD	28-JUN-2001.	
PF	22-DEC-2000; 2000WO-US035395.	
XX		
PR	23-DEC-1999; 99US-0173003P.	
PA	(CORR.) CORNELL RES FOUND INC.	
XX	(NISC-) JAPAN SCI & TECHNOLOGY CORP.	
PI	Goldman SA, Okano H;	
DR	WPI; 2001-418053/44.	
XX		
PT	Separating multipotential neural progenitor cells from a mixed population	
XX	of cells, involves introducing nucleic acid molecule encoding fluorescent	
XX	protein under promoter control, and separating fluorescent cells.	
PS	Claim 12, Fig 11A-UJ; 87pp; English.	
XX		
CC	The present sequence represents the human muasashi promoter, which is used	
CC	in the method of the invention. The specification describes a method for	
CC	separating multipotential neural progenitor cells from a mixed population	
CC	of cell types. The method comprises introducing a nucleic acid encoding a	
CC	fluorescent protein under control of a promoter which selectively	
CC	functions in the progenitor cells, into all cell types of the mixed	
CC	population, allowing only the progenitor cells to express the fluorescent	
CC	protein, and identifying and separating cells that are fluorescent, where	
CC	the separated cells are progenitor cells. The method is useful for	
CC	isolation and purification of multipotential neural progenitor cells,	
CC	especially neural stem cells from adult brain. The isolated cells are	
CC	used in both basic analyses of precursor and stem cell growth control, as	
CC	well as in more applied studies of their transplantability and	
CC	engraftment characteristics. The cells are useful in support of the	
CC	structural repair of the damaged central nervous system, such as in the	
CC	traumatized brain, or the contoured, traumatized or transected spinal	
CC	cord	
XX		
SQ	Sequence 52216 BP; 13494 A; 12461 C; 12362 G; 13899 T; 0 U; 0 Other;	
Query Match	12.6%; Score 252.6; DB 4; Length 52216;	
Best Local Similarity	69.1%; Pred. No. 4, 8e-47;	
Matches 403; Conservative	0; Mismatches 174; Indels 6; Gaps 4;	
DB		
417	TTTTCACATTTTTTAAGTAGCACTTAATAACAGGTAAAAAAGCGTGCCGACAGTGGCTCAG	476
41515	TTAAAAATTTTCTTAGGAACAACACTTAAA--AGACATTAAAGGCCGCGGAGCTCAC	41572
477	ACCTGATATCCCAAGCACTTTGGAGGCTGAGGCAGGCAAGTCACTTTGGTCAGAGTTT	536
41573	TTCGTGATATCCCAAGCACTTTGGAGGCGCAGAGGCAAGTGATCATCTGAAGTCAAGAGTTG	41632
537	GAGACTGCTGCTGGCCAACATGCGAAACTCTGCTCTTAAAAAAAATTACAAATTTAGCC	596
41633	GAGACCGAGCCTGGGCAACAGAGGTAAACCATGTCTCT-ACATAAAATTAACAAATTTAGCT	41691

Query Match	12.6%	Score 252.4	DB 11	Length 56098
Best Local Similarity	70.8%	Pred. No. 5.5e-47		
Matches 408	Conservative 0	Mismatches 156	Indels 12	Gaps 5
01-MAR-2002; 2002US-00087192.				
(SAGR-) SAGRES DISCOVERY.				
Morris DW;				
WPI: 2003-328604/31.				
Claim 1; SEQ ID NO 1612; 0pp; English.				
The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published				
Sequence 56098 BP; 17842 A; 10579 C; 10322 G; 16942 T; 0 U; 413 Other;				
440 TTAAACAGGTAAGGCTGGGCGAGTGGCTCACCTGTAATCCAGACACTTTGGG				499
26680 TTAAAACTGATGACAGGCCAGGCGGTGGCTCACACTGTAAATCCAGAGTTGGG				26221
500 AGGTGAGGCGAGGAGATCACTTTGGTCAAGATTGAGATAGCTGGCCAACTGGC				559
26620 AGGTGAGGCGGAGATCACTTGAAGTTGAAACAGAGCTGGCCAACTGGT				26561
560 GAACTCTGTCTTAAAAAATAAATTAGCTGGCATGTGGCGGGGCGCTGTA				619
26560 GAAACCCCATCTT-ACATAAAATAGAAAGTTAGCCGATGTGTGGCACACGCTTGT				26502
620 ATCTCAGCTGTCAAGAGGCGGAGACACAAGATCACTTAAACCCAGAGGTGGAGTTG				679
26501 GTCCCAACGCTCGGAGGCGAGAGGCCAAGATCACTTGACCCCAAGTGGAGTTG				26442
680 CAGTACGCTGAGATCGTCCCA-CTCACTCCACCTGGGAGACAGAGTGCACCTTTGTCT				738
26441 CAGTACGCTGAGATTGAGTACGCACTCCAGCTGGGAGACAGAGATCTCACTTC				26382
739 CAAAAAGAAAAAATAAATAAAGAAAAAAGAGTGAAGTTAACTTAAATACCAA				798
26381 AGAAAAATAAATAA-----CAAAACAAAACTGATATGTTAATTTTAAATAATA				26328
799 TGTATCCCAATATCAATCACTTCAAGTGTAAATTAATATTAACAATATATGAGAT				858
26327 TTTATTTTAAACCGTTATATCTAAAAATGCTCATTTGACATGTATCAATGTGA				26268
859 CTTTACATCTCTTTCTTGTGTTTCA-TTAAAGCTTTGAAGAAGAGATATATATGATG				917
26267 ATGGAATATTTTACATCTTTTAAATAATAGCTTTTGAAGTCCAGAGTGTATATAC				26208
918 TGACAGACACA---TCTCAATTTGAGATAGCTATCACTTCAAGTGTCAAGTACCACTGTG				974
26207 TTTATAGCAACAATCTTCAATTTTGAATAGCAATCTTCAAGTGTCTCAACAGGACATGTG				26148
975 GCTAGCAGTTACTGTATTTGATGGACGAGATCTTGA 1010				

Db	26147	GCCAGTGGCTACTGTTGGATPAACATATTCTAGA	26112
RESULT	8		
ID	ACA60949/C		
AC	ACA60949 standard; DNA; 55827 BP.		
XX	ACA60949;		
DT	11-AUG-2003 (first entry)		
XX	DNA encoding human carboxypeptidase.		
DE			
XX			
KM	Human; gene; carboxypeptidase; inflammation; cancer; arteriosclerosis;		
KW	neurodegenerative disease; protease; db; single nucleotide polymorphism;		
XX	SNP.		
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	replace(858,T)		
FT	/*tag= u		
FT	/note= "Single nucleotide polymorphism"		
FT	2017. .53409		
FT	/*tag= a		
FT	/product= "Carboxypeptidase"		
FT	2017. .2090		
FT	/*tag= b		
FT	/number= 1		
FT	2091. .19206		
FT	/*tag= c		
FT	/number= 1		
FT	replace(2122,T)		
FT	/*tag= v		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(4088,A)		
FT	/*tag= w		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(4240,A)		
FT	/*tag= x		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(4416. .4418,TA)		
FT	/*tag= y		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(4434,G)		
FT	/*tag= z		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(4965,C)		
FT	/*tag= aa		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(5868,A)		
FT	/*tag= ab		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(6254,C)		
FT	/*tag= ac		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(10171,C)		
FT	/*tag= ad		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(10452,A)		
FT	/*tag= ae		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(11613,A)		
FT	/*tag= af		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(12130,T)		
FT	/*tag= ag		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(17867. .17869,GA)		
FT	/*tag= ah		
FT	/note= "Single nucleotide polymorphism"		
FT	replace(18243,G)		
FT	/*tag= ai		

```
FT /note= "Single nucleotide polymorphism"
FT exon 19207. .19282
FT /tag= d
FT /number= 2
FT 19283. .22683
FT /tag= e
FT /number= 2
FT replace(22450,C)
FT /tag= aj
FT /note= "Single nucleotide polymorphism"
FT exon 22684. .22808
FT /tag= f
FT /number= 3
FT 22809. .24477
FT /tag= g
FT /number= 3
FT replace(23003,T)
FT /tag= ak
FT /note= "Single nucleotide polymorphism"
FT variation replace(24055. .24057,GA)
FT /tag= al
FT /note= "Single nucleotide polymorphism"
FT variation replace(24132,T)
FT /tag= am
FT /note= "Single nucleotide polymorphism"
FT exon 24478. .24586
FT /tag= h
FT /number= 4
FT 24587. .28129
FT /tag= i
FT /number= 4
FT replace(25713,A)
FT /tag= an
FT /note= "Single nucleotide polymorphism"
FT exon 28130. .28231
FT /tag= j
FT /number= 5
FT 28232. .33052
FT /tag= k
FT /number= 5
FT replace(28476,C)
FT /tag= ao
FT /note= "Single nucleotide polymorphism"
FT variation replace(29404,G)
FT /tag= ap
FT /note= "Single nucleotide polymorphism"
FT variation replace(31047,T)
FT /tag= aq
FT /note= "Single nucleotide polymorphism"
FT variation replace(31445,C)
FT /tag= ar
FT /note= "Single nucleotide polymorphism"
FT variation replace(31447,C)
FT /tag= as
FT /note= "Single nucleotide polymorphism"
FT variation replace(31600,C)
FT /tag= at
FT /note= "Single nucleotide polymorphism"
FT variation replace(31714,C)
FT /tag= au
FT /note= "Single nucleotide polymorphism"
FT variation replace(31715,T)
FT /tag= av
FT /note= "Single nucleotide polymorphism"
FT variation replace(32193,G)
FT /tag= aw
FT /note= "Single nucleotide polymorphism"
FT variation replace(32341,A)
FT /tag= ay
FT /note= "Single nucleotide polymorphism"
FT variation replace(32341,G)
FT /tag= ax
FT /note= "Single nucleotide polymorphism"
```

```
FT variation replace(32561,G)
FT /tag= az
FT /note= "Single nucleotide polymorphism"
FT variation replace(32600. .32602,AA)
FT /tag= ba
FT /note= "Single nucleotide polymorphism"
FT variation replace(32642,T)
FT /tag= bb
FT /note= "Single nucleotide polymorphism"
FT variation replace(32793. .32795,TG)
FT /tag= bc
FT /note= "Single nucleotide polymorphism"
FT exon 33053. .33157
FT /tag= 1
FT /number= 6
FT replace(33071,A)
FT /tag= bd
FT /note= "Single nucleotide polymorphism"
FT intron 33158. .42288
FT /tag= m
FT /number= 6
FT replace(34721,T)
FT /tag= be
FT /note= "Single nucleotide polymorphism"
FT variation replace(35304,A)
FT /tag= bf
FT /note= "Single nucleotide polymorphism"
FT variation replace(35425,C)
FT /tag= bg
FT /note= "Single nucleotide polymorphism"
FT variation replace(36050,G)
FT /tag= bh
FT /note= "Single nucleotide polymorphism"
FT variation replace(36291,G)
FT /tag= bi
FT /note= "Single nucleotide polymorphism"
FT variation 42289. .42382
FT /tag= n
FT /number= 7
FT 42383. .48643
FT /tag= o
FT /number= 7
FT exon 48644. .48846
FT /tag= p
FT /number= 8
FT 48847. .51176
FT /tag= q
FT /number= 8
FT exon 51177. .51264
FT /tag= r
FT /number= 9
FT 51265. .53224
FT /tag= s
FT /number= 9
FT exon 53225. .53406
FT /tag= t
FT /number= 10
FT XX
PN US2003017574-A1.

Query Match 12.0%; Score 240.2; DB 8; Length 55827;
Best Local Similarity 68.6%; Pred. No. 3.3e-44;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTTAGAGTGTGTGTTTGTAGAAATATAGCTGAGCTGCTTATGTCATTTCGA 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46943 TGAATCCACGACGCGGTTGTTATAGAAATATATAGCAGTCACATATTACTTTAA 46884
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 423 GTTTTGTAGCACCATTAAACAGTAAAC--AAGGCTGGCGCAGTGCTCACACT 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 46883 ATTTCAGTTCACCTTTAAGAAAATTAAGAAAGAGCGACGCGTGCTCACACT 46824
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 481 GTAATCCAGCATTGGAGGCTGAGCGACGACATCATTGTGTCAGAGTTTGAGA 540
```


46823 ATATCCACACCTTTGGAGGCTGAGGAGCGGATCATCTTGATCAGAGTTTACA 46764
46823 ATATCCACACCTTTGGAGGCTGAGGAGCGGATCATCTTGATCAGAGTTTACA 46764
QY 541 CTAGCTGGCCCAACATGCGGAAACTGTGTCTCTAAAAAATACAAAATAGCCGGC 600
QY 46763 CCAGCGAGGCCAACATGCTGTAACCCCGCTCTACCAAAA--TACAAAATAGCCAGGC 46706
QY 601 ATAGTGGCGGCGGCTTAACTTCAGCTGCTCAGAGAGCCGAGACAGAAATCATTAA 660
QY 46705 ATAGTGGTGAACGCTGTAAATCCAGCTACTTGGAGGCTGAGGACAGAACTTGCTTGA 46646
QY 661 ACCGAGAGGTGAGGTGAGTTCAGTGAATGCTGCTCACT--CACTCAACTGGGAGA 719
QY 46645 ACCGAGAGGACAGAGGTGAGTGAATGCTGAGATCATGACTGCACTCCAGCCGTGATGA 46586
QY 720 CAGAGTGACA-CTTTGTCTCAAAAAAGAAAAAACAATGAAAAAGAACAGGTGA 778
QY 46585 CAGAAACAGACCCCTGCACACACACACAAAGAAATATGTGAGATTATTCATA 46526
QY 779 AGTTAACTTAATACCAATGTATCCCAATACAAATCATTTCAAGTGAATTAATA 838
QY 46525 TTGTAAATTTAATTAACCAAAATGTACAAAGTATCA---TTCAAGTGAACCAATATA 46470
QY 839 AAACAATTATGATGATGATCTTACATCTTTCTGTCTTCAATTAAGTCTTTGAAA 898
QY 46469 AA--AATTATCTTGAATAATTTTATGTACTT-----TTTCACTAAGCTTCAAT 46417
QY 899 GTATGATATATATGTTATGCTGACAGACATCTCAATTTGCACTGACTACATTTGAGTGC 958
QY 46416 TCAGTGTGATTAATTAACATCAGTGAATGTCTCAATTCAGACAGATCAATTTTACAC 46357
QY 959 TCAGTACCCACATGTGCTGACATCTACTGATTTGA 995
QY 46356 TGGATAGCCACATGTGTTAGTGTGCTACTGATGCTGA 46320

RESULT 9

ABX13671/c
ID ABX13671 standard; DNA; 55827 BP.

XX ABX13671;

DT 14-FEB-2003 (first entry)

XX Human protease gene.

XX Human; gene; ds; protease; proteolytic degradation; proteolysis;

KM proliferation; differentiation; signaling; therapeutic; gene therapy;

KM protein therapy; diagnostic; immune response; vaccine; inflammation;

KM cancer; arteriosclerosis; degenerative disorder; chromosome 13;

XX single nucleotide polymorphism; SNP.

XX Homo sapiens.

XX Key location/Qualifiers

XX variation replace(858,T)

XX CDS /tag= u

XX /standard_name= "Single nucleotide polymorphism"

XX /tag= a

XX /product= "Protease"

XX /tag= b

XX /number= 1

XX /tag= c

XX /number= 1

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= x

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= y

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= z

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= aa

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ab

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ac

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ad

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ae

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= af

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ag

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ah

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ai

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= aj

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ak

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= al

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= am

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= an

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ao

FT /standard_name= "Single nucleotide polymorphism"

FT /tag= ap

FT /standard_name= "Single nucleotide polymorphism"

```
FT exon 28130..28231
FT /*tag= j
FT /number= 5
FT 28232..33052
FT /*tag= k
FT /number= 5
FT replace(28476,C)
FT /tag= ao
FT /strand.name= "Single nucleotide polymorphism"
FT replace(29404,G)
FT /*tag= ap
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31047,T)
FT /*tag= aq
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31445,C)
FT /*tag= ar
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31447,C)
FT /*tag= ab
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31600,C)
FT /*tag= ac
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31714,C)
FT /*tag= au
FT /strand.name= "Single nucleotide polymorphism"
FT replace(31715,T)
FT /*tag= av
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32193,G)
FT /*tag= aw
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32341,R)
FT /*tag= ax
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32561,G)
FT /*tag= ay
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32600..35601,AAA)
FT /*tag= az
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32642,T)
FT /*tag= ba
FT /strand.name= "Single nucleotide polymorphism"
FT replace(32793..32794,TAG)
FT /*tag= bb
FT /strand.name= "Single nucleotide polymorphism"
FT 33053..33157
FT /*tag= 1
FT /number= 6
FT replace(33071,A)
FT /*tag= bc
FT /strand.name= "Single nucleotide polymorphism"
FT 33158..42288
FT /*tag= m
FT /number= 6
FT /cons.splice= (5'site:no,3'site:no)
FT replace(34721,T)
FT /*tag= bd
FT /strand.name= "Single nucleotide polymorphism"
FT replace(35304,A)
FT /*tag= be
FT /strand.name= "Single nucleotide polymorphism"
FT replace(35425,C)
FT /*tag= bf
FT /strand.name= "Single nucleotide polymorphism"
FT replace(36050,G)
FT /*tag= bg
FT /strand.name= "Single nucleotide polymorphism"
FT replace(36291,G)
FT /*tag= Dn
FT /strand.name= "Single nucleotide polymorphism"
```

```
FT exon 42289..42382
FT /*tag= n
FT /number= 7
FT 42383..48643
FT /*tag= o
FT /number= 7
FT /cons.splice= (5'site:yes,3'site:no)
FT 48644..48846
FT /*tag= p
FT /number= 8
FT 48847..51176
FT /*tag= q
FT /number= 8
FT /cons.splice= (5'site:yes,3'site:no)
FT 51177..51264
FT /*tag= r
FT /number= 9
FT 51265..53224
FT /*tag= 8
FT intron
```

Query Match 12.0%; Score 240.2; DB 10; Length 55827;

Best Local Similarity 68.6%; Pred. No. 3.3e-44; Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

```
Qy 363 TCATACCTAGAGTAGTGCTGTTAGTAGAAATTAATGCTGAGCTGCTTATGTCATTTC 422
Db 46943 TGAATCCTCAGACAGGGTGTGTTAATAGAAATTAATGCAATGTCATATTACTTTAA 46884
Qy 423 GTTTTATTAGTACCATTTAAACAGGTAAA--AAGCTGGCCGACGTGCTCACCT 480
Db 46883 ATTTTCAGTTACCACTTTAAAGAAAATTAATAAGGACAGGACGGTGGCTCACACT 46824
Qy 481 GTATCCAGCACTTTGGAGGCTGAGCAGCAGATCACTTTGGTCAGAGTTTGAGA 540
Db 46823 ATATATCCAGCACTTTGGAGGCTGAGCAGGCGATCACTTGAATCAGAGTTTGAGA 46764
Qy 541 CTAGCTGGCCAAATGCGAAACTCTGCTTAATAAATAATACAAATTAAGCTGCG 600
Db 46763 CAGGCGAGGCCAAATGCGAAACCCGCTCTACCAAAA--TACAAATAATGCGAGGC 46706
Qy 601 ATGATGGCGGGGCGCTGTAATCTCAGCTGCTCAGAGGCGGACACAAAGATTCCTTA 660
Db 46705 ATAGTGTGCACGCGCTGTAATCCAGCTACCTGGAGGCTGAGGACGAACTGCTTGA 46646
Qy 661 ACCAGAGAGTGAAGTTSCAGTGAAGTGAATCGGCACCT--CACTCCAACTGGAGAG 719
Db 46645 ACCAGAGAGCAGAGGTTGCAAGTGAAGTGAATCAGACTGCACTCCAGCTCGATGA 46586
Qy 720 CAGAGTGACA--CTTTGTCTCAAAAAGAAAAAACAAGTMAAAAAAGAAACAGGTGA 778
Db 46585 CAGAACAGACCTGCGACACACACACAAAAAAGAAATGTGAGATTAAATTCATTA 46526
Qy 779 AGTTAATTAAATTAACCAATGATATCCCAATCAATCAATTTCAAGTGTAAATTAATA 838
Db 46525 TTGTAAATTAATTAACCAATGATCAAAAGATCA---TTTCAAGTGAACCAATTA 46470
Qy 839 AAACATTAATGAATGATTAATCTTACATCTTTTCTGTTCAATTAAGCTTTGAAA 898
Db 46469 AA--AATTAATCTTGAATAATTTTATGTTACTT-----TTTCATATCAATAGCTTAAAT 46417
Qy 899 GTGAGTATATATGTTATGCTGACAGCAATCTCAATTTGGAAGTGAATTCAGGTGC 958
Db 46416 TCTGGTGTATTTAACTCAAGTGAATGTCTCAATTCAGACAGGTCATTTTAAGCAC 46357
Qy 959 TCAGTAGCCACATGTGGCTAGCAGATTACTGTATTGGA 995
Db 46356 TGGATGCCACATGTGTTAGTGCGCTACTATCTGGA 46320
```

RESULT 10

ADS36454 standard; DNA; 58337 BP.

ID ADS36454

XX

CC genetic polymorphisms associated with an altered risk of developing an
CC autoimmune disease (e.g. rheumatoid arthritis). The invention further
CC comprises a method of identifying an individual that has an altered risk
CC of developing an autoimmune disease, comprising detecting a single
CC nucleotide polymorphism (SNP) in a nucleic acid of the invention. The DNA
CC and protein sequences of the invention are useful for diagnosing and
CC treating autoimmune diseases, such as: rheumatoid arthritis, type 1
CC diabetes, multiple sclerosis, systemic lupus erythematosus, inflammatory
CC bowel diseases, psoriasis, thyroiditis, celiac disease, pernicious
CC anemia, asthma, vitiligo, glomerulonephritis, Grave's disease, the
CC myocarditis, Sjogren's disease, or primary systemic vasculitis. The
CC present nucleic acid represents a human autoimmune disease-related
CC genomic DNA sequence of the invention. NOTE: The present sequence is not
CC shown in the specification, but has been retrieved from the WIPO website.

XX
XX
SQ Sequence 64423 BP; 19548 A; 12552 C; 12463 G; 19737 T; 0 U; 123 Other;

Query Match 12.0%; Score 240.2; DB 13; Length 64423;

Best Local Similarity 68.6%; Pred. No. 3.4e-44;
Matches 437; Conservative 0; Mismatches 183; Indels 17; Gaps 7;

QY 363 TCATACCTCTAGTGTGTTTGTAGTAAATATGCTAGCTGCTATGCTATTCCTCA 422
DB 51306 TGAATCTCTGACACAGGCTGTGTTAATGAAATATGCAAGTCACATATTTACTTTAA 51247
QY 423 GTTTTATAGTACCATTTAAACAGGTAA--AAGGCTGGCGCAGTGGCTCACACCT 480
DB 51246 ATTTTCAAGTTACCACTTTAAGAAAATATAAGAAAGGCGACGCTGCTCAACCT 51187
QY 481 GTATATCCACACTTTGGAGGCTGAGGCGACGATCCTTTGTCAGAGTTTGAGA 540
DB 51186 ATATATCCACACTTTGGAGGCTGAGGCGGATCATCTGAGATCAGAGTTTGACA 51127
QY 541 CTAGCTGGCCACATGCGGAAACTGTCTCTAAAAAATATACAAAATTGCGCTGGC 600
DB 51126 CAGGCGAGGCCAACATGCTAAACCCGCTCTACCAAAAA--TACAAAAATGCGCAGGC 51069
QY 601 ATGCTGGCGGCGCTGTATCTCAGCTGCTCAGAGGCGGACACAAATATCATTTAA 660
DB 51068 ATAGTGTGACCCGCTGTATCCAGTACTTTGGAGGCTGAGGCGAATCTTGCTTGA 51009
QY 661 ACCGAGAGGTGAGAGTTGTCAGTGTGATGCTGATGCTCACT--CACTCCAACTGGGAGA 719
DB 51008 ACCGAGAGGTGAGAGTTGTCAGTGTGATGCTGATGCTCACTCCAGCTGGATGA 50949
QY 720 CAGAGTGACA--CTTTTGTCTCAAAAAGAAAAAACAAGTAAATAAAGAAACAGGTGA 778
DB 50948 CAGAAACAAGCCCTGCCACACACACACAAAAAATATGTGATTTAATTCATA 50889
QY 779 AGTTAATTTAATTAACCAATGATCCCAATATCAATCATTTCAAGTGTAAATTAATA 838
DB 50888 TTGTAATTTAATTAACCAATATGTAACAAGTATCA----TTTCAAGTGTAAACAATATA 50833
QY 839 AAAACATTAATGATGATATCTTTAATCTCTTTCTTTGTTTCAATTAAGCTTTGAAA 898
DB 50832 AA--AATTAATCTTGAATATTTTAAATTTTAACTT-----TTTCAATTAAGCTTTCAAT 50780
QY 899 GTAGATATATATGTTATGCTGAGACGACATCTCAATTTGAGTACAGTCACTTCAGGTGC 958
DB 50779 TCTGGGTGTATTTAAACATCACTGATGTTGTTCAATTCAGACAAAGTACATTTTAAGCAC 50720
QY 959 TCAGTACGACATGTGGCTAGCAGTTACTGTAATTTGA 995
DB 50719 TGGATAGCCACATGTGGTATGCTATGCTGGA 50683

RESULT 12

ACN45066/c
ID ACN45066 standard; DNA; 27189 BP.

XX ACN45066;

DT 18-NOV-2004 (first entry)

XX
XX
DE Human genomic sequence hCG310694.
XX
XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

XX Homo sapiens.

XX MO2003073826-A2.

XX 12-SEP-2003.

XX 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1828; Opp; English.

XX
XX
CC The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

XX
SQ Sequence 27189 BP; 6410 A; 6780 C; 6720 G; 7279 T; 0 U; 0 Other;

Query Match 11.8%; Score 235.6; DB 11; Length 27189;

Best Local Similarity 68.8%; Pred. No. 3e-43;
Matches 385; Conservative 0; Mismatches 164; Indels 11; Gaps 4;

QY 439 ATTAAAAACAGTAAATAAGCTGGCGCAGTGGCTCAGCTGTATATCCAGACCTTTGG 498
DB 11830 AGTAAATATGCTAAATAATGCGCGGCGATGCTCATCTGTATATCCACACTATGA 11771
QY 499 GAGGCTGAGGCGAGGACATCACTTTGTCAGAGGTTGAGACTGAGCTGGCGCAACATGG 558
DB 11770 GAGGCGAGGAGGCGAGTCACTTGAAGTCAGAGGTTTGAACACAGCTGGCGCAATAG 11711
QY 559 CGAAATCTGTCTCTTAAAAAATAAATAAATTAAGCTGGAGTGGCGGCGCTGT 618
DB 11710 TGAATCCCGCTCTCT--ACTAAAGATATTAATTAAGTGGAGTGGTGGGTGCTCT 11652
QY 619 AATTCAGTGTCTCAGAGGCGGAGACACAGAAATCATTAACCCAGAGGTGAGGTT 678
DB 11651 AATCCAGTACTTCAAGAGGCTGAGGAGAGAAATTTGTAACCCAGAGGAGGAGTT 11592
QY 679 GAGAGAGGTGAGATGCTGCCACTCACTCCAACTGGGAGAGAGATGACATTTTGTCT 738
DB 11591 GAGAGAGCGAGAGTCACTCACTTCAGCTCCAGATGAGTGCAGAGGCGAAAGTCTGT 11532
QY 739 CAAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 798
DB 11531 CAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11477
QY 799 TGTATCCAAATPACATCTTTCAAGTGTAAATTAATTAATAAATAAATAAATAAATAAATAA 858

Db	11476	TACACCTCCCTCCCTTCATTTTCAAAAAAATTTTCATGCAATTAAACAAACATTATTAA	11417
Qy	859	CTTTACACTTCTTTCTTGTGTTTCATATTAAGCTTTGAAAGAGATATATGTTATGCG	917
Db	11416	GATTACACTTC-----TCTCTCATGTGTAAGCTTTGAAATCTGGGGTATATTTATACT	11361
Qy	918	TGACAGACACTTCATTTTGGACTAGTATACATTTCAAGTGCTCAAGACCAATGTGGCT	977
Db	11360	TAAAGGATGCTCTTAAATTTGGACTAGACACATTTCAAGTGCTCAGTAGCTACACGTGAT	113010
Qy	978	AGCAGTACTGTATTGGATG	997
Db	11300	AGTGCTCACTGTATTTAGATG	11281

RESULT 13
ACM43998_3/c
Constitution (4 of 7) of ACM43998 from base 300001 (Human genomic sequence hCG1778483.)
MP Sequence split into 7 fragments LOCUS ACM43998 Accession ACM43998
ID Database Name Ref

Fragment Name	Begin	End
ACN43998_0	1	11000
WB		21000
ACN43998_1	10001	11000
WB	20001	31000
ACN43998_2	30001	41000
WB	40001	51000
ACN43998_3	50001	61000
WB	60001	653122
ACN43998_4		
ACN43998_5		
ACN43998_6		

Query Match	11.5%	Score 229.2;	DB 11;	Length 110000;
Best Local Similarity	69.7%;	Pred. No. 1.3e-41;		
Matches 404; Conservative	0;	Mismatches 153;	Indels 23;	Gaps 6;

QY	440	TTAAACAGGTTAAAAAGGCTGGGGGAGGAGGCTCAACCTGTAAATCCAGCACTTTGGG	499
Db	60810	TTTTAAAAATACAAATTGGGCGAGGTGGGGGCTCAATGTGTAAATCCAGCACTTTGAA	60751
QY	500	AGGTTGAGGACGAGCATCACTTTGGTCAGAGTTTGAGACTAGGCTGGCCAAATGCG	559
Db	60750	AGGCGGAGGGGGGGAATCAC--GAGGTACAGAGTTTGAGACCAAGCTGACCAATGGT	60693
QY	560	GAAACTGTGTCTTAAAAAAAATACAAAATTAGCTGGCATGTGGCGGGCGCTGTGA	619
Db	60692	GAAACCCCGGCTCT-ACATAAAATTACAAAATTAGCGAGGCGTGGTGTGCAAGCTGTGA	60634
QY	620	ATCTCAGCTCTCAGAGGCGCGAGACACAAGATTCATTAACCCAGAGAGTGGAGGTTG	679
Db	60633	ATCCCAACCACTTGGGAGGCTGTGAGACAGGGAATCACTTGATCTTGGAGGCGAGAGTTG	60574
QY	680	CAGTGAAGCTAGATTCGTGCACAT-CAGTCCAACTGGGAGACAGATGACACTTTGTCT	738
Db	60573	CAATGAATGAGATTATGCCACTGTCACTCCAGCTGGGGCGACAGACCAATCTCCATCTG	60514
QY	739	CAAAAAGAAAAAAGAAAAACAAGTAAAAAGAAAAGGTGAATTAACTTTAATACCNA	798
Db	60513	AAAAAAAAAAAAAAAAAAAAAAAAAAAAATTTGG-----ATTTACTTAA	60466
QY	799	TGATCCCAATACAAATCATTTTCAAAAGTAAATTAATATTAACATTAATGAGATGATA	858
Db	60465	TATATCCAAAATGTTTAAATTTCAACATGTCACTTAATATTAAGAAATGAGATGATTA	60406
QY	859	CTTTCACATCTTTTCTGTTTTCATATTAAGTCTTTGAAGAGATATATGTTATGCT	918
Db	60405	ACCTTCTTCTTCTTTTTC---TACTAAGCTCTTGGATGCCAGCATATTTTATACA	60349
QY	919	GACAGCAATCTCAATTGGACTAGCT---ACATTTCAAGGTCCTGATAGCCACATGTG	974
Db	60348	TATGGCTATCTTGATTTGGGGCTAGCTAGCCAGCTTCCAGGGCTCAAGGCCACATCTG	60289
QY	975	GCTAGCAGTTACTGTATTGATGGCACCGGATCTTAGAGGA	1014
Db	60288	GCTGTGTAGCTACCATVAGCAAGATGACGAGACTTAGAAGCA	60249

RESULT 14	
AAK68359/c	
ID	AAK68359 standard; DNA; 2953 BP.
XX	
AC	AAK68359;
XX	
DT	06-NOV-2001 (first entry)
XX	
DE	Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23171.
XX	
KW	Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM	cytostatic; gene therapy; vaccine; metastasis; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200157182-A2.
XX	
PD	09-AUG-2001.
XX	
PF	17-JAN-2001; 2001WO-US0001354.
XX	
PR	31-JAN-2000; 2000US-0179065P.
PR	04-FEB-2000; 2000US-0180628P.
PR	24-FEB-2000; 2000US-0184664P.
PR	02-MAR-2000; 2000US-0186350P.
PR	16-MAR-2000; 2000US-0189874P.
PR	17-MAR-2000; 2000US-0190076P.
PR	18-APR-2000; 2000US-0198123P.
PR	19-MAY-2000; 2000US-0205515P.
PR	07-JUN-2000; 2000US-0209467P.
PR	28-JUN-2000; 2000US-0214886P.
PR	30-JUN-2000; 2000US-0215135P.
PR	07-JUL-2000; 2000US-0216647P.
PR	07-JUL-2000; 2000US-0216880P.
PR	11-JUL-2000; 2000US-0217487P.
PR	11-JUL-2000; 2000US-0217487P.
PR	14-JUL-2000; 2000US-0218230P.
PR	26-JUL-2000; 2000US-0220963P.
PR	26-JUL-2000; 2000US-0220964P.
PR	14-AUG-2000; 2000US-0224518P.
PR	14-AUG-2000; 2000US-0224519P.
PR	14-AUG-2000; 2000US-0225213P.
PR	14-AUG-2000; 2000US-0225214P.
PR	14-AUG-2000; 2000US-0225266P.
PR	14-AUG-2000; 2000US-0225267P.
PR	14-AUG-2000; 2000US-0225268P.
PR	14-AUG-2000; 2000US-0225270P.
PR	14-AUG-2000; 2000US-0225447P.
PR	14-AUG-2000; 2000US-0225477P.
PR	14-AUG-2000; 2000US-0225757P.
PR	14-AUG-2000; 2000US-0225758P.
PR	14-AUG-2000; 2000US-0225759P.
PR	18-AUG-2000; 2000US-0226279P.
PR	22-AUG-2000; 2000US-0226681P.
PR	22-AUG-2000; 2000US-0226688P.
PR	22-AUG-2000; 2000US-0227182P.
PR	23-AUG-2000; 2000US-0227009P.
PR	30-AUG-2000; 2000US-0228924P.
PR	01-SEP-2000; 2000US-0229287P.
PR	01-SEP-2000; 2000US-0229343P.
PR	01-SEP-2000; 2000US-0229344P.
PR	01-SEP-2000; 2000US-0229345P.
PR	05-SEP-2000; 2000US-0229509P.
PR	05-SEP-2000; 2000US-0229513P.
PR	06-SEP-2000; 2000US-0230437P.
PR	06-SEP-2000; 2000US-0230438P.
PR	08-SEP-2000; 2000US-0231242P.
PR	08-SEP-2000; 2000US-0231243P.
PR	08-SEP-2000; 2000US-0231244P.
PR	08-SEP-2000; 2000US-0231413P.
PR	08-SEP-2000; 2000US-0231414P.
PR	08-SEP-2000; 2000US-0232080P.
PR	08-SEP-2000; 2000US-0232081P.
PR	12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234978P.
PR 25-SEP-2000; 2000US-0234988P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236328P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241866P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.

PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 23171; 3071bp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;
SQ
Query Match 11.4%; Score 227.6; DB 4; Length 2953;
Best Local Similarity 69.0%; Pred. No. 1,1e-41;
Matches 421; Conservative 0; Mismatches 164; Indels 25; Gaps 7;
QY 414 TCATTTCCAGTTTATTTAGTACCATTTAAACAGGTAAACAA-----GCTGGCGCA 467
DB 1699 TAATTAAAGATTTTAAACAGCCACATTAATGAAGAAACAAAGTCGGCGGCGG 1640
QY 468 GTGGCTCAACCTGTAATCCGACCTTTGGAGGTGAGGAGGATCACTTTGGT 527
DB 1639 GTGGCTCAACCTGTAATCCGACCTTTGGAGGTGAGGAGGATCACTTTGGT 1582
QY 528 CAGAGATTGAGATCTGACCTGGCCAAATGCGAACTGTCTCTAAACAAATATCA 587
DB 1581 CCGAAGATGAGATCTGACCTTTGTAACAGCGTAACTCCGCTCTCAATTAATCA 1522
QY 588 AAATTAGCTGACATGTCGCGCGCTGTAATCTCACTCTCTCAGAGCGCGAGAC 647
DB 1521 AAATTAGCTGACATGTCGCGCGCTGTAATCTCACTCTCTCAGAGCGCGAG 1462
QY 648 AAGATCACTTAAACCCAGAGGTGAGGTGACATGAGTCTGCTCACTCACTC 707
DB 1461 GAGAAATGCGGTAAACCCGAGGAGGAGCTTCAAGTCACTGATGAGCACTGCACT 1402
QY 708 CAACCTGGAGAGCAGAGTACACTTTTCTCAAAAAAGAAAAAACAAGTAAAA 767
DB 1401 CCAACGCTGGGCGAGAGAGC-AGACTCTGTCTCAAAAAAAGAAAAAAGAAAA 1343
QY 768 GAAACAGGTGAAGTTAATTTAATTAACCAATGTATCCAAATACATCTTCAAA 827

Db 1342 TAAGTCATAGATTTAATATATTAGCTAGATATGCTAAATATATCATTTAA----- 1288
Qy 828 TAATATATATAAACATTTATGATGAGATCTTACTTCTTTCTTTGTTTATATTA 887
Db 1287 -CAGTATATCAAAAGATTATTAACAAATTTTACATTCCTGTTGG---GGGTATATA 1232
Qy 888 AGCTTTGAAAGAGATATATA-TGTTATGCTGACAGACATCTCAATTTGGACTAGCTA 946
Db 1231 AATCTTTGAATTAATATGTTATTTTATTTACTTACAGCTGATCTTGATTCAGACTAGCCA 1112
Qy 947 CATTTCAGTGTCTCAGTAGCCATGTGGCTAGCA-----GTTACTGTATTTGATGCA 1000
Db 1171 CATTTCAGTGTCTCAGTAGCCCATGTGGCTAGCAAGATGTTACTATATGATATGCA 1112
Qy 1001 CGGATCTTGA 1010
Db 1111 CAGCTTATTA 1102

RESULT 15
AAK68360/C
ID AAK68360 standard; DNA; 2953 BP.
XX AAK68360;
AC AAK68360;
XX
DT 06-NOV-2001 (first entry)
XX
XX Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:23172.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
XX cytotoxic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205151P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 11-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225447P.
XX
XX 14-AUG-2000; 2000US-0225757P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 16-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226681P.
XX
XX 22-AUG-2000; 2000US-0226686P.

PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246539P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.

17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 23172; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I), proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention
XX
XX Sequence 2953 BP; 860 A; 572 C; 619 G; 902 T; 0 U; 0 Other;

Query Match 11.4%; Score 227.6; DB 4; Length 2953;
Best Local Similarity 69.0%; Pred. No. 1.1e-41;
Matches 421; Conservative 0; Mismatches 164; Indels 25; Gaps 7;

QY 414 TCATTTCCAGTTTTTTAGTACCATTTAAACAGGTAAAAA-----GGCTGGCGCA 467
DB 1699 TAATTAAGATTTTTTTAAGCACAATTATATATGTAAGAAAAAGTCGGCGGCGG 1640
QY 468 GTGGCTCACACCTGTATATCCAGCATTTGGAGGCTGAGGCAAGGCATACCTTTGGT 527
DB 1639 GTGGCTCACACCTGTATATCCAGCATTTGGAGGAGACCGAGGCGGATCAC--GACGT 1582

QY 528 CAGGATTTGAGACTAGCTGCGCAACATGCGAACTGCTCTTAATAAAAAATATCA 587
DB 1581 CCGAGATCGAACCATCTTGTTAACAGCGGAAATCCGCTCTACTATAAAAATACAA 1522
QY 588 AAATTAGCTGCGCATGTGCGGCGCCTGTATCTCAGCTGCTCAGAGGCGGAGAC 647
DB 1521 AATTAGCCAGGCATGTGCGGATGCTGTAGTCCAGCTACTTGGAGGCTGAGGCG 1462
QY 648 AAGATCACTTAAACCCAGAGGTGAGGTGCACTGAGCTGAGATCGGCCACTCATC 707
DB 1461 GAGAAATGGGTAAACCCGAGGAGAGAGCTTCAGTGAAGTGAAGCAGCATGCACT 1402
QY 708 CAACCTGGAGACAGAGTGACACTTTTGTCAAAAAAGAAAAAACAAGTAAAAA 767
DB 1401 CACCCAGCTGGGCGCAGAGAC -AGACTCTGCTCAGAAAAAAGAAAAAAGAAAAA 1343
QY 768 GAAACAGGTGAAGTTAATTATATACCAATGTATCCAAATACATCATTTCAAGTG 827
DB 1342 TAAATCATGATTTTATATATTTAGCTCAGATGTCTAAATATATCATTTAA----- 1288
QY 828 TAATTAATTAATAAATATATGAGATTAATTACATTTCTTCTTGTTCATATTA 887
DB 1287 -CATGTATCAACAAAGTTATTAACAAATATTTCAATTCCTGTGG--GGTTAATA 1232
QY 888 AGTCTTGAAGTGAATATATA-TGTTATGCTGACAGACATCTCAATTGGAAGTATGA 946
DB 1231 AATCTTGAATTAATAATGTATTTTATTAATTAACAGCTGATCTTGATTAAGCTAG 1172
QY 947 CATTCAGGTGCTCAGTACCAATGTGCTAGCA-----GTACGTATTTGATGGA 1000
DB 1171 CATTCAGGTGCTCAGTACCCCAATGTGCTAGCAATGTGTTACTATATGAATAGCA 1112
QY 1001 CGATCTAGA 1010
DB 1111 CAGCTCTATA 1102

Search completed: February 25, 2005, 20:15:00
Job time : 1024.11 secs

forms: with and without stop-codon (to allow fusion with C-terminal tag). The CDS has been directionally cloned using BD In-Fusion(TM) cloning system between the Sali and HindIII sites of the pDNR-DUAL vector. Additional sequences in the clone: 'ACC' after Sali site and before 'ATG' to provide Kozak consensus sequence; 'GG' after last codon and before HindIII site to maintain reading frame. Clone distribution: <http://bioinfo.clontech.com/orfclones>.

FEATURES

source

```

1..882
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="GH00284X1.0"
/clone_lib="BD Creator (TM) CDS Library derived from MGC collection"
/lab_host="DH5alpha TI resistant"
/note="Vector: pDNR-Dual"

```

CDS

```

1..882
/codon_start=1
/product="kallikrein 5"
/protein_id="PAP35513.1"
/db_xref="GI:30582573"
/translation="MATARPMMWVLCALITALLVTEHVLANDVSCDHSNTVPS
GNSQD/GAAGEEDARSDDSSRLINGSDCDMTQPMQALLRPNOLYCAVLYHPQW
LTPAHCRRKRVFRLVGLHYSLSPEVSGQMGVKSIPHGVSHLGHNSNDMLIKLN
RRIKPTKDVPIVNSHSPGASAGTKVSGMGTKSPQVHPKVLQCLNLSVLSORKE
DAVPRQIDDTMPCAGKAGRDSQGGSGGVVGNGLQGLVNSGDYPCARPNRPGVYT
NICKTKMIOETIQANS"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      7..98e-121      Length:      882
Score:          1608.00        Matches:      293
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    100.00%        Indels:        0
DB:             9              Gaps:           0

```

US-09-936-271B-14 (1-293) x BT006867 (1-882)

```

QY      1 MetAlaThrAlaArgProPOTPTpMetTPRValLeuCYaAlaLeuIleThrAlaLeu 20
DB      1 ATGGCTAAGACGACAGACCCCTGATGAGTGGTCTCTGCTCATCAGAGCTTGCTT 60
QY      21 LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerArg 40
DB      61 CTGGGGGTCAACAGCATGTTCTCGCCACATGATGTTCTCTGTGACCACTCCCTTAC 120
QY      41 ThrValProSerGlySerAspGlnAspLeuGlyValGlyValGluAspAlaArgSer 60
DB      121 ACCGTGCTCTGTGGAGCAACGAGCACTGGAGCTGGGGCCGGGGAGACCGCCGGTGG 180
QY      61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMethIsthArgInProTP 80
DB      181 GATGACAGACGACGCGCATCATCATGATGATCCGACTGCGATATGACACCCGCGCTGG 240
QY      81 GlnAlaAlaLeuLeuLeuArgProAspGlnLeuTyrCysGlyValValLeuValHisPro 100
DB      241 CAGGCGCGCTGTGCTTACAGGCCCAACACAGCTTACTGCGGGCGGGTGTGGTCAATCCA 300
QY      101 GlnTPRLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
DB      301 CAGTGGCTGCTCAACGCGCGCCACTGCAAGAGAAAGATTTCAGAGTCCGTCTCGGCCAC 360
QY      121 TyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValIleValSerIle 140
DB      361 TACTCCCTGTCAACACATTAATGATCTGGGACACAGATGTTCCAGGGGGTCAATCATC 420
QY      141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleValSerIle 160
DB      421 CCCACCTCGGTACTCCACCTCGGCACTCTTAAGCACTCAAGCTCATCAAACTGAGAC 480
QY      161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180

```

```

DB      481 AGAAGATTGTCCTCCACTTAAGATGCAAGCCATCAAGCTCTCTCATGTCCCTCT 540
QY      181 AlaGlyThrIleCysLeuValSerGlyTTPGlyThrThrIleSerProGlnValHisPhe 200
DB      541 GCTGGGACAAATGCTGTGTGTCTGTGCTGGGGGACAAACCAAGACCCCAAGTGCATTC 600
QY      201 ProValIleuGlnCysLeuAsnIleSerValIleuSerGlnIleArgCysGluAspAla 220
DB      601 CTTAAGTCTCTCAAGTCTTGAATATACCGCTCTTAAGTCAGAAAAGTGCAGATGCT 660
QY      221 TTPRArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValIleArgAspSer 240
DB      661 TACCCGACAGATGATGATACACCATGTTCTGCGCGCGGAGCAAAAGCAGTAGAGCTCC 720
QY      241 CysGlnGlyAspSerGlyIleProValValCysAsnGlySerLeuGlnIleValSer 260
DB      721 TGCACAGGTGATTCGGGGGGCTGTGTGCTTCATAGCTCCCTGACGAGACTCGTTC 780
QY      261 TTPGlyAspTyrProCysAlaArgProAspArgProGlyValTyrThrAsnLeuCysIle 280
DB      781 TGGGAGATTAACCTTGTCTGCTCCGCGCCACACAGCCGGGTGTCTACAGAACTTCTCAAG 840
QY      281 PheThrIleTTPIleGlnIleThrIleGlnAlaAsnSer 293
DB      841 TTCACCAATGATTCAGAGAAACCATCCAGGCCAATCC 879

```

RESULT 2

BT007831

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

JOURNAL

AUTHORS

COMMENT

CDS

source

1..882

/organism="synthetic construct"

/mol_type="mRNA"

/db_xref="taxon:32630"

/clone="GH00284X1.0"

/clone_lib="BD Creator (TM) CDS Library derived from MGC collection"

/lab_host="DH5alpha TI resistant"

/note="Vector: pDNR-Dual"

1..882

/note="Mutations: 881:Stop->Leu"

DB: 9 Gaps: 0
US-09-936-271b-14 (1-293) x AY279380 (1-1370)
QY 1 MetAlaThrAlaArgProProTrrMetTrrValLeuCyAlaLeuLeuLeu 20
DB 128 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGCTGTGATCAAGCTTGGCT 187
QY 21 LeuG1VValThrG1uH1eValLeuAlaAsnAspValSerCyAspHisProSerAen 40
DB 188 CTGGGGGTACAGAGATGTTCTCCGCAACATGATGTTCTGTATCAACCCCTCTAAC 247
QY 41 ThrValProSerG1ySerAsnG1nAspLeuG1yAlaG1yAlaG1yAlaAspAargSer 60
DB 248 ACCGTCCCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAGAGCCCGGTG 307
QY 61 AspAspSerSerSerArg1le1leAsnG1ySerAspCyAspMetHisThrG1nProTrr 80
DB 308 GATGACAGACGAGCGGATCATCATGATCGATCGATATGACACCCAGCCGTG 367
QY 81 G1nAlaAlaLeuLeuLeuArgProAsnG1nLeuTrrCyG1yAlaValLeuValHisPro 100
DB 368 CAGGCCGCTGTGTGTAAGGCCCAACAGCTCTACTGGGGGGGTGTGTGATCCA 427
QY 101 G1nTrrPLeuLeuThrAlaAlaHisCyAsArg1yValPheArgValArgLeuG1yHis 120
DB 428 CAGTGGCTGTCAACGCGCCGCTGCAAGAGAAAGTTTTCAGATCCGTCTCGGCAC 487
QY 121 TrrSerLeuSerProValTrrG1uSerG1yG1nMetPheG1nG1yVallylsSer1le 140
DB 488 TACTCCCTGTACACAGATTATGATCTGGGAGAGATGTTCCAGGGGTCAATCCATC 547
QY 141 ProHisProG1yTrrSerHisProG1yHisSerAsnAspLeuMetLeu1leu1leu1leu 160
DB 548 CCCCACCTGGCTACTCCACCTGGCACTCTTAACGACTCATGCTCATCAACTGAAC 607
QY 161 ArgArg1leArgProThrTrrAspValArgPro1leAsnValSerSerHisCyProSer 180
DB 608 AGAAGATTCGTCCCACTAAGATGTCAAGCCCATCAAGCTCTCTCATTTGCTCTCT 667
QY 181 AlaG1yThrLyCyAlaLeuValSerG1yTrrG1yThrThrLySerProG1nValHisPhe 200
DB 668 GCTGGGACAAAGTGTGTGTCTGGCTGGGGGACCAACAGAGCCCAAGGCACTTC 727
QY 201 ProLyValLeuG1nCySerLeuAsn1leSerValLeuSerG1nLyAsArgCyG1uAspAla 220
DB 728 CTTAAGCTCTCCAGTGTCTTAATATCAAGCTCTTAAGTCAAGAAAGGTGCGAGATGCT 787
QY 221 TrrProArgG1n1leAspAspThrMetPheCyAlaG1yAspLyAlaG1yArgAspSer 240
DB 788 TACCCGAGACAGATAGATGACACCATGTTCTGGCGCGGTGACAAACAGATGAGACTCC 847
QY 241 CyG1nG1yAspSerG1yG1yProValValCyAsnG1ySerLeuG1nG1yLeuValSer 260
DB 848 TCCCAAGGTATCTGGGGGGCTGTGTGTCTGCAATGGCTCTCTGAGAGGACTCGTCTCC 907
QY 261 TrrPGLyAspTrrProCyAlaArgProAsnArgProG1yValTrrThrAsnLeuCyLyAs 280
DB 908 TGGGAGATTAACCTTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACTCTGACAG 967
QY 281 PheThrLySer1leG1nG1yThr1leG1nAlaAsnSer 293
DB 968 TTCACCAAGTGAATCCAGAAACCATCCAGGCCCACTCC 1006
RESULT 4
BD107879 1381 bp DNA linear PAT 18-SBP-2002
LOCUS BD107879 36 human secreted proteins.
DEFINITION BD107879
ACCESSION BD107879
VERSION BD107879.1 GI:23202697
KEYWORDS JP 2002500035-A/50.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1381)
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002500035-A 50 08-JAN-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002500035-A/50
PD 08-JAN-2002
PR 06-JAN-1999 JP 2000527554
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN
NI,
PI LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
C12N15/09,A61K31/711,A61K38/00,A61K39/395,A61K39/395,A61K48/00,
PC A61P5/00,
PC A61P7/00,A61P11/06,A61P19/02,A61P29/00,A61P31/18,A61P35/00, PC
A61P35/02,
PC A61P37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC
C12N15/00,
PC A61K37/02,C12N5/00
PC 36 human secreted proteins
FH Key location/Qualifiers
FT source 1..1381
FT location/Qualifiers
FEATURES
source 1..1381
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1,296-120 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-936-271b-14 (1-293) x BD107879 (1-1381)
QY 1 MetAlaThrAlaArgProProTrrMetTrrValLeuCyAlaLeuLeuLeu 20
DB 129 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGCTGTGATCAAGCTTGGCT 188
QY 21 LeuG1VValThrG1uH1eValLeuAlaAsnAspValSerCyAspHisProSerAen 40
DB 189 CTGGGGGTACAGAGATGTTCTCCGCAACATGATGTTTCTGTGACCAACCCCTCTAAC 248
QY 41 ThrValProSerG1ySerAsnG1nAspLeuG1yAlaG1yAlaG1yAlaAspAargSer 60
DB 249 ACCGTCCCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAGAGACCCCGGTG 308
QY 61 AspAspSerSerSerArg1le1leAsnG1ySerAspCyAspMetHisThrG1nProTrr 80
DB 309 GATGACAGACGAGCGGATCATCATGATCGATCGATATGACACCCAGCCGTG 368
QY 81 G1nAlaAlaLeuLeuLeuArgProAsnG1nLeuTrrCyG1yAlaValLeuValHisPro 100
DB 369 CAGGCCGCTGTGTGTAAGGCCCAACAGCTCTACTGGGGGGGTGTGTGATCCA 428
QY 101 G1nTrrPLeuLeuThrAlaAlaHisCyAsArg1yValPheArgValArgLeuG1yHis 120
DB 429 CAGTGGCTGTCAACGCGCCGCTGCAAGAGAAAGTTTTCAGATCCGTCTCGGCAC 488
QY 121 TrrSerLeuSerProValTrrG1uSerG1yG1nMetPheG1nG1yVallylsSer1le 140
DB 489 TACTCCCTGTACACAGATTATGATCTGGGAGAGATGTTCCAGGGGTCAATCCATC 548

QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleuValLeuAsn	160
Db	549	CCCCACCCTGGGTACTCCACCCTGGCACTTAACGACTATGCTATCAACTGAAC	608
QY	161	ArgArgIleArgProThrIleAsnValArgProIleAsnValSerSerHisCysProSer	180
Db	609	AGAAATAATTCGTCCCACTAAAGATGTCAGACCCTCAACAGTCTCTCATATGTCTCCCT	668
QY	181	AlaGlyThrIleCysLeuValSerGlyTyrTrpGlyThrThrIleValSerProGlnValHisPhe	200
Db	669	GCTGGAGCAAAATGCTTGCTGTGCTCTGGCGGGGACAACCAAGACCCCAAGTGAATTC	728
QY	201	ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleArgCysGluAspAla	220
Db	729	CCTAAGTCTCTCAAGTGTGTAATATCAAGCGTGCTAAGTCAGAAAAGGCGCAGGATGCT	788
QY	221	TyrProAlaGlnIleAspAspThrMetPheCysAlaGlyAspIleValGlyValArgAspSer	240
Db	789	TACCCGAGACAGATGATGATGACACCATGTTCTGCGCGGTGAACAAAGCAGGTAGAACATCC	848
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
Db	849	TGCCAGGGGTGANTCTGGGGGGCTGTGGCTTCGCAATGGCTCCCTCGAAGGACTCGTGTCC	908
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysAllys	280
Db	909	TGGGAGATTACCTCTTGCTGCCCCGCGCCAAACAGACGGGGTGTTCACAGAACTCTGCAAG	968
QY	281	PheThrIleTyrIleGlnIleThrIleGlnIleAsnSer	293
Db	969	TTTACCAATGTGATCCAGAAACCATTCAGAGCCAATCC	1007

LOCUS	CO874884	1387 bp	DNA	linear	PAT 27-SEP-2004
DEFINITION	Sequence 2 from Patent WO2004077060.				
ACCESSION	CO874884				
VERSION	CO874884.1	GI:52748034			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Euteria; Primates; Carnivora; Homnidae; Homo.				
JOURNAL	1 Diamond, E.P. and Petraki, C.				
FEATURES	Assay for detection of renal cell carcinoma				
Source	Patent: WO 2004077060-A 2 10-SEP-2004;				
	Mount Sinai Hospital (CA)				
	Location/Qualifiers				
	1..1387				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Alignment Scores:					
Pred. No.:	1,296-120	Length:	1387		
Score:	1608.00	Matches:	293		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US-09-936-271B-14 (1-293) x CO874884 (1-1387)					
Qy	1 MetalaatrraiaarproptroptmrttrpvalleucysalaleuilethrAlaleu	20			
Db	170 ATGCTACAGCAAGACCCCTCGATGCGGCTCTGAGCTCATACAGCCTGCTT	229			
Qy	21 leuglvalththgluhisvalleulalahnmaharvalsercysaaphisproserasn	40			
Db	230 CTGGGGGTACAGACACATGTTCTGGCCACAACATGATGTTCTCTGACCACCCCTTAC	289			

QY	41	ThrValProSerGIYseAenGIaAspLeuGIyLaGIyLaGIyLaAspAlaPArSer	60
Db	290	ACCGTGCCCTTGGAGACACCAAGACTTGGAGACTGGGCCGGGAGAGAGCCCGGTGC	349
QY	61	AaPArSeSerSeSerArGIaLeIleAaNGIYseArPcYBaSPMeChIstHGIaProTrP	80
Db	350	GATGACAGCAGCAGCCGGATCATCAATGGATTCGAGTCGGATATGCAcACCAGCCGGTGG	409
QY	81	GIaLaAlaLeuLeuLeuArgProAenGIaLeuTYrCysGIyAlaValLeuValHisPro	100
Db	410	CAGCCCGGGCTGTGTCAAGGCCCAACAGCTCTACCGCGGGGGGTGGTGATCA	469
QY	101	GIaTrPLeuLeuThrAlaAlaHisCYaArgYbAlaYsValPheaArgValArgLeuGIYHis	120
Db	470	CAGTGCTGTCTCAAGCGCGCCCACTGACAGAGAAAGTTTCAAGATCCGTCTCGGCCAC	529
QY	121	TYrSerLeuSeProValTYrGIuSeGIyGIaNGIaMetPheGIaGIyValYsSerIle	140
Db	530	TACTCCCTGCACAGATTATGATCTGGGACAGATGGTTCCAGGGGGTCAATTCATCC	589
QY	141	ProHIsProGIYTYrSerHisProGIYHisSerAaAspLeuMetLeuIleYsLeuAa	160
Db	590	CCCCACCCTGGCTACTCCCAcCTGGGCACCTCTAACACCTCAAGCTCATCAAACTGAAC	649
QY	161	ArgArgIleArgProThrLYsAspValArgProIleAaValSerSerHisCYaProSer	180
Db	650	AGAAGATTCGTCCCACTTAAGAAGTCAGACCCATCAACGCTCTCTCATATGTCCTCT	709
QY	181	AlaGIyThrLYsCYsLeuValSerGIYTrpGIYThrThrLYsSerProGIaValHisPhe	200
Db	710	GCTGGAGCAAAAGTCTTGGTGCTGGCGTGGGGGCAACCAAGACCCCCCAAGTCACTTC	769
QY	201	ProLYsValLeuGIaCYsLeuAaHisSerValLeuSerGIaLYsArgCYsGIuAspAla	220
Db	770	CCTAAGGCTCCCAcGTCTTGAATATACGCTGTCAAGTCAGAAGAGGTGGAGAGATGCT	829
QY	221	TYrProArgGlnIleAaPArThMetPheCYaAlaGIyAspLYsValaGIYArgAAsPArSer	240
Db	830	TACCCGAGACAGATAGATGACACCATGTTCTGCCCGGTGACAAAGCGATAGAGACTCC	889
QY	241	CysGIaNGIYAsPSeGIyGIYProValCYaAaNGIYseTrLeuGIaNGIYLeuValSer	260
Db	890	TGCCAGGGTGAATTCTGGGGGGCGTGTGTCGAATGGCTCCCTGACGAGGACTGTGTCC	949
QY	261	TrpGIYAspTYrProCYsAlaArgProAaArgProGIYValTYrThrAsnLeuCYaLYs	280
Db	950	TGGGAGATTACCTTGTGTGCCCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAG	1009
QY	281	PheThrLYsTrpIleGIaNGIaThrIleGIaAlaAaNSer	293
Db	1010	TTTCCCAAGTGTGATCTCAGAAACCATCAAGGCCAATCTCC	1048

RESULT 6					
LOCUS	CQ874960	1387 bp	DNA	linear	PAT 27-SEP-2004
DEFINITION	Sequence 4 from Patent WO2004075713.				
ACCESSION	CQ874960				
VERSION	CQ874960.1	GI:52748059			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Diamandis, E.P.				
TITLE	Multiple marker assay for detection of ovarian cancer				
JOURNAL	Patent: WO 2004075713-A 4 10-SEP-2004;				
	MOUNT SINAI HOSPITAL CORPORATION (CA)				
FEATURES	location/Qualifiers				
source	1. 1387				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:

Pred. No.: 1 296-120 Length: 1387
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x C0874960 (1-1387)

QY 1 MetAlaThAlaArProProTriPMeCTrPValLeuCYaAlaLeuIleThAlaLeuLeu 20
 Db 170 ATGGCTACGAGCAAGACCCCTGGATGTGGGTCTGTGCTCTGATACAGCTTGCTT 229
 QY 21 LeuGlyValThrGluHisValLeuAlaAaAaAaPValSerCYaAaPHisProSerAaA 40
 Db 230 CTGGGGGTCAAGAGCATGTTCTGCCAACAATGATTTCTGTGACCAACCCCTTAAAC 289
 QY 41 ThrValProSerGlySerAaAaAaPLeuGlyValAGlyValAGlyValAaAaAaAaAa 60
 Db 290 ACCGTGCTCTGGAGCAACGAGCTGGAGCTGGAGCCGGAGAGAGCCCGGCTGG 349
 QY 61 AaPAsPSeSerSerAa 80
 Db 350 GATGACAGCAGCAGCCGATCATCATGATGATCGACTGGCATATGCACACCCAGCCGTG 409
 QY 81 GlnAlaAlaLeuLeuLeuAa 100
 Db 410 CAGGCGGCTGTGTGTAAAGCCCAACAGCTTACTGGGGGGGGTGTGGTGTCTTCA 469
 QY 101 GlnTrpLeuLeuThAlaAlaHisCYaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 120
 Db 470 CAGTGGCTGTCTACGCGCCGCCCACTGCAAGAAAGATTTCAGATCCGTCTCGCCAC 529
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValYSerIle 140
 Db 530 TACTCCCTGTCAACAGTTATGATCTGGGCAAGATGTTCCAGGGGGTCAATCATC 589
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAaAaAaAaAaAaAaAaAaAaAaAaAa 160
 Db 590 CCCACCTGGCTACTCCACCCCTGGCACTTAAAGACTTCAATGCTCAAACTGAAC 649
 QY 161 ArgArgIleArgProThrIleAaPValAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 180
 Db 650 AGAAGATTCGCTCCACTAAAGATGCAAGCCATCAAGCTCTCTCTCATTTGCTCTCT 709
 QY 181 AlaGlyThrIleCYaLeuValSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
 Db 710 GCTGGGACAAAGTCTTGGGTGTGGCTGGGGGCAACCAAGAGCCCCCAAGTGCATTC 769
 QY 201 ProLeuValLeuGlnCYaLeuAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 220
 Db 770 CCTAAAGTCTCCAGCTTGAATATCAAGGTCTTAAAGTCAAGAAAGTGCAGAGTGTCT 829
 QY 221 TyrProArgGlnIleAaPAsPThrMetPheCYaAlaGlyAaPValAGlyValArgAaP 240
 Db 830 TACCCAGACAGATGATGACACCATGTTCTGCGCGGAGCAAGAGAGTAAAGACTCC 889
 QY 241 CYaGlnGlyAaPSeSerGlyProValValCYaAaAaAaAaAaAaAaAaAaAaAaAaAa 260
 Db 890 TGCAGAGGTGATCTGGGGGGCTGTGTGCTGCAATGGGTCCCTGCAAGGAGCTCGTCTC 949
 QY 261 TrpGlyAaPValProCYaAa 280
 Db 950 TGGGAGATTAACCTTGTGCTGCGGCCAAGACAGACCGAGTGTCTTACAGAACTCTCAAG 1009
 QY 281 PheThrIleValIleGlnGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 293
 Db 1010 TTCACCAAGTGAATCAAGAAACCATTCAGGCCAATCTCC 1048

RESULT 7

AF168768 1387 bp mRNA linear PRI 16-APR-2002
 LOCUS Homo sapiens stratum corneum tryptic enzyme (SCTE) mRNA, complete
 DEFINITION cds
 ACCESSION AF168768
 VERSION AF168768.2 GI:20153423
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE 1 (bases 1 to 1387)
 purification, molecular cloning, and expression of a human stratum

JOURNAL J. Biol. Chem. 274 (42), 30033-30040 (1999)
 MEDLINE 99445563
 PUBMED 10514489

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUL-1999) Public Health and Clinical Medicine,

REFERENCE
 AUTHORS Brattand, M. and Egelrud, T.
 TITLE 3 (bases 1 to 1387)
 Dermatology and Venereology, Umea University, University Hospital,

JOURNAL Submitted (16-APR-2002) Public Health and Clinical Medicine,
 Dermatology and Venereology, Umea University, University Hospital,
 Umea SE-901 85, Sweden

REMARK
 COMMENT Sequence update by submitter
 On Apr 16, 2002 this sequence version replaced gi:6063032.

FEATURES
 source location/Qualifiers
 1..1387
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="stratum corneum"
 1..1387
 /gene="SCTE"
 170..1051
 /gene="SCTE"
 /note="SCTE, trypsin-like serine protease; possibly
 involved in desquamation"
 /codon_start=1
 /product="stratum corneum tryptic enzyme"
 /protein_id="AA03101.1"
 /db_xref="GI:6063033"
 /translation="MATARPMPWVLCALITLIGVTEHYLVANNVSCDHSNTVPS
 GSNODLAGAGEDARDSDSSRIINGSDCDHTOPQMAALLRPQLYCGAVLVHPOW
 LITLAHCKRVPRVLGHYSLSPVESGQMGQVKSIPHPGSHSGNSNDMLTKLN
 RIRPRKXVRPNTVNSHCPASCTKLVSGMKVTKSPQVHFPRVLOCLINISVSORCE
 DAVPRQIDITMECADDKGRDSCQDGSQGVVCSLQGLVSMGDPICARPNRPVYT
 NICKETKWIQETIQANS"

ORIGIN

Alignment Scores:

Pred. No.: 1 296-120 Length: 1387
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x AF168768 (1-1387)

QY 1 MetAlaThAlaArProProTriPMeCTrPValLeuCYaAlaLeuIleThAlaLeuLeu 20
 Db 170 ATGGCTACGAGCAAGACCCCTGGATGTGGGTCTGTGCTCTGATACAGCTTGCTT 229
 QY 21 LeuGlyValThrGluHisValLeuAlaAaAaAaPValSerCYaAaPHisProSerAaA 40

Db 230 CTGGGGGTCAAGAGCATGTTCTCGCCAAAGATGATGTTCTCTGACACCCCTCTAAC 289
Qy 41 ThrvAlProSerGlySerAenGlnAapLeuGlyValAglValAglValAapPalaAarSer 60
Db 290 ACCGTGCTCTTGGAGCAACCAAGACCTGGGAGCTGGGGCGGGAAGAGCCGGTGG 349
Qy 61 AspAapSerSerSerAarGilelleAasnGlySerAapCyAapAmetHsthGlnProTrp 80
Db 350 GATGACAGCAGCAGCCGATCATATGATGATCCAGCTGCATATGACACCCAGCCGTGG 409
Qy 81 GlnAlaAlaLeuLeuLeuAarGProAenGlnLeuTyrcyGlyValAvalLeuValHsPro 100
Db 410 CAGGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGGGGGGGGGTGGTGTGATCCA 469
Qy 101 GlnTrpLeuLeuThrAlaAlaHsCyAarGlyValValPheAarGlyValArgLeuGlyHis 120
Db 470 CAGTGGCTGTCTACAGGGCGGCGGCTGACAGGAAGAAAGTTTCAAGATCCGCTCGGGCAC 529
Qy 121 TyrserLeuSerProValTyrgLusSerGlyGlnGlnMetPheGlnGlyValLysSerTle 140
Db 530 TACTCCCTGTCAACAGTTTATGATCTGGGCGAGAGATGTTCCAGGGGGTCAATTCATC 589
Qy 141 ProHsAProGlyTyrserHsAProGlyHsSerAapAapLeuMetLeuLysLeuAasn 160
Db 590 CCCACCTGGCTACTCCACCTGGGCACTGTAACGACCTCATGCTCATCAAACTGAC 649
Qy 161 ArhArgGlnLeaRProThrlYsAapValArgProLleAenValSerSerHsCyAProSer 180
Db 650 AGAAGATTCGTCTCCACTTAAAGATGACAGCCCTCAACGTCTCTCATATGCTCTCT 709
Qy 181 AlAglYThrLysCyALeuValSerGlyTrpGlyThrTrpLysSerProGlnValHsPhe 200
Db 710 GCTGGGCAAAAGTCTTGCTGTCTGGCTGGGGGCAACCAAGACCCCAAGTCACTTC 769
Qy 201 ProLysValLeuGlnCyALeuAasnLysSerValLeuSerGlnYsArgCyGluAapPala 220
Db 770 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGCTGTAAGTGAAGAAAGTGGAGAGAGCT 829
Qy 221 TyrTrpAArgGlnLleAapAapTrpMetPheCyAalAglYsAlyValAglYsAapSer 240
Db 830 TACCCGAGACAGATGATGACACCATGTTCTGCCCGGTGCAAAAGAGTAGAGACTCC 889
Qy 241 CyseGlnGlyAapSerGlyYsProValValCyAasnGlySerLeuGlnGlyLeuValSer 260
Db 890 TGCAGGGGTATCTGGGGGGCTGTGTCTGCAATGACTCTCCAGGAGACTGTCTTC 949
Qy 261 TrpGlyAapTyrcyCyAalAarGProAasnAarGProGlyValTyrrHsAenLeuCyAlys 280
Db 950 TGGGAGATTAACCTTGTGCGGCGCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1009
Qy 281 PheThrLysTrpLleGlnGlyThrTrpLleGlnAlaAasnSer 293
Db 1010 TTCACCAAGTGAATCCAGAAACCATCAGGCCCAACTCC 1048

RESULT 8
AY279381 1438 bp mRNA linear PRI 26-MAY-2003
LOCUS Homo sapiens kallikrein 5 splice variant 2 (KLK5) mRNA, complete
DEFINITION cds; alternatively spliced.
ACCESSION AY279381
VERSION AY279381.1 GI:31075482
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1438)
AUTHORS Kurlender, L., Yousef, G.M., White, N.M.A., Robb, J.-D., Borgono, C.A.
and Diamandis, E.P.
TITLE Identification of splice variants for the human kallikrein gene 5
(KLK5)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1438)

AUTHORS Yousef, G.M., Robb, J.-D., White, N.M.A., Kurlender, L., Borgono, C.A.
and Diamandis, E.P.
TITLE Direct Submission
JOURNAL Submitted (19-Apr-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada

FEATURES
source
1. 1438
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. 1438
/gene="KLK5"
1. 195
/gene="KLK5"
196. 1077
/gene="KLK5"
/note="serine protease; differentially expressed in
malignancy; may play a role in desquamation of skin; hks;
alternatively spliced; similar to the product of Genbank
Accession Number AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 2"
/protein_id="AAP42276.1"
/db_xref="GI:31075483"
/translation="MATARPMMWVLCALITALLGVTEHTVLANDVSCDPSNTVPS
LSNQDAGAGGEBARSDSSRIINGSDCMHPQALILRPNOLYCAVIVHPQM
LTPAARCKKVPFRLGHSLSPIVSSGQMPQGVKSIPIHGYSHPSHMDLILKN
RRIPTKDVPIVSSHCPAGKICLVSGMTKSPVHPVLYLQCLNISVLSQKCE
DAYRQIDPTMPCAGDKAGRDSQGSQSGGVVNGSLQGLVSGDYPCARPNNRPGVYT
NICKFTKWIOETIQANS"

3' UTR
polyA_signal
1078. 1438
/gene="KLK5"
1387. 1392
/gene="KLK5"

ORIGIN
Alignment Scores:
Pred. No.: 1,34e-120 Length: 1438
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x AY279381 (1-1438)

Qy 1 MetAlaThrAlaArgProPOTtrpMetTrpValLeuCyAAlaLeuLleThrAlaLeuLeu 20
Db 196 ATGGCTACAGCAAGACCCCTGATGTGGTCTCTGTGCTTGAATCAGAGCTTGGCTT 255
Qy 21 LeuGlyValThrGlnHsValLeuAlaAasnAapValSerCyAapHsAProSerAasn 40
Db 256 CTGGGGGTCAAGAGCATGTTCTGCGCCACAAAGATGTTTCTGTGACCAACCCCTCTAAC 315
Qy 41 ThrvAlProSerGlySerAenGlnAapLeuGlyValAglValAglValAapPalaAarSer 60
Db 316 ACCGTGCTCTTGGAGCAACCAAGACCTGGAGCTGGGGCGGGAAGAGCCGGTGG 375
Qy 61 AspAapSerSerSerAarGilelleAasnGlySerAapCyAapAmetHsthGlnProTrp 80
Db 376 GATGACAGCAGCAGCCGATCATATGATGATCCAGCTGCATATGACACCCAGCCGTGG 435
Qy 81 GlnAlaAlaLeuLeuLeuAarGProAenGlnLeuTyrcyGlyValAvalLeuValHsPro 100
Db 436 CAGGCGCGCTGTGTCTAAGGCCCAACAGCTTACTGGGGGGGGGTGGTGTGATCCA 495
Qy 101 GlnTrpLeuLeuThrAlaAlaHsCyAarGlyValValPheAarGlyValArgLeuGlyHis 120
Db 496 CAGTGGCTGTCTACAGGGCGGCGGCTGACAGGAAGAAAGTTTCAAGATCCGCTCGGGCAC 555

QY 121 TYSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIlySerSile 140
|
|
|
Db 556 TACTCCCTGTCACAGATTATGATCTGGGACAGCAATGTTCCAGGGGGGCAAAATCATC 615
QY 141 ProH1sProGlyTyrSerH1sProGlyH1sSerAsnAspLeuMetLeuIleIlyLeuAsn 160
|
|
|
Db 616 CCCACCCCTGGCTACCTCCACCCCTGGCCACTCTAACGACCTCATGTCTCAAACTGAAC 675
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerH1sCysProSer 180
|
|
|
Db 676 AGAAGATTCCGTCCCACTTAAGATGTCAGACCATCAAGCTCTCTCATTTGCTCCCTCT 735
QY 181 AlaGlyThrIlyCysLeuValSerGlyTyrTPGlyThrThrIlySerProGlnValH1sPhe 200
|
|
|
Db 736 GCTGGGACAAAGCTTGGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCACTTC 795
QY 201 ProlValIleuGlnIlyLeuAsnIleSerValLeuSerGlnIlyArgCysGluAspAla 220
|
|
|
Db 796 CCTAAGTCTCTCCAGCTTGAATATCACGCTGAAGTCAGAAAGGTGCGAGATGCT 855
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyValArgAspSer 240
|
|
|
Db 856 TACCCAGACAGATGATGACACCACTGTTCTGGCCGGTGACAAAGCAAGTAGAGCTCC 915
QY 241 CysGlnGlyAspSerGlyGlyProValIlyCysAsnGlySerLeuGlnIlyLeuValSer 260
|
|
|
Db 916 TGCACAGGTGATCTGGGGGGCTGTGTCTGCTGCAATGGCTCCCTGCAGGGGACTCGTCTC 975
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlyS 280
|
|
|
Db 976 TGGGAGATTACCTTGTGCTCCCGGCCCAACAGACCGGGTGTCTACGAACTCTGCAG 1035
QY 281 PheThrIlySTrpIleGlnIlyuThrIleGlnIlyAsnSer 293
|
|
|
Db 1036 TTCACCAAGTGATCCAGGAACCATCCAGGCCCACTCC 1074
RESULT 9
AR352504 1499 bp DNA linear PART 17-AUG-2003
LOCUS Sequence 1 from patent US 6589770.
DEFINITION AR352504
ACCESSION AR352504
VERSION AR352504.1 GI:33757710
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1499)
AUTHORS Kitado,H., Yoshikawa,A. and Zaiki,T.
TITLE Keratinocyte derived protease
JOURNAL Patent: US 6589770-A 1 08-JUL-2003;
FEATURES Location/Qualifiers
source 1..1499
/organism="unknown"
/mol_type="genomic DNA"
ALIGNMENT SCORES:
Pred. No.: 1..4e-120 Length: 1499
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-936-271b-14 (1-293) x AR352504 (1-1499)

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValIlyGluAspAlaArgSer 60
|
|
|
Db 411 ACCGTGCTCTTGTGGAGGACCAAGACCTGGAGACTGGGGCCGGGGAAGACCCCGGTG 470
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetH1sThrGlnProTP 80
|
|
|
Db 471 GATGACAGCAGACCGGCATCATGATGATCCGACTGGATGATACACCCAGCCGTGG 530
QY 81 GlnAlaAlaLeuLeuAsnArgProAsnGlnLeuTyrCysGlyValIlyValIlyAspAla 100
|
|
|
Db 531 CAGGCGGCTGTCTGTAAGGCCCAACAGCTTACTGCGGGCCGGTGTGGATGCCA 590
QY 101 GlnTPLeuLeuThrAlaAlaH1sCysArgIlySlyValPheArgValArgLeuGlyH1s 120
|
|
|
Db 591 CAGTGGCTGCTACCGGCCGCCACGACGAGGAAGAAAGTTTCAAGTCCGTCTCGGCCAC 650
QY 121 TYSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIlySerSile 140
|
|
|
Db 651 TACTCCCTGTCACAGTTTATGATCTGGGACAGATGTTTCCAGGGGGTCAAAATCCATC 710
QY 141 ProH1sProGlyTyrSerH1sProGlyH1sSerAsnAspLeuMetLeuIleIlyLeuAsn 160
|
|
|
Db 711 CCCACCCCTGGCTACTCCACCTGACCTTAACGACCTCATGCTCATCAAACTGAAC 770
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerH1sCysProSer 180
|
|
|
Db 771 AGAAGATTCCGTCCCACTTAAGATGTCAGACCATCAAGCTCTCTCATTTGCTCCCTT 830
QY 181 AlaGlyThrIlyCysLeuValSerGlyTyrTPGlyThrThrIlySerProGlnValH1sPhe 200
|
|
|
Db 831 GCTGGGACAAAGTGTGTGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCACTTC 890
QY 201 ProlValIleuGlnIlyLeuAsnIleSerValLeuSerGlnIlyArgCysGluAspAla 220
|
|
|
Db 891 CCTAAGTCTCTCAGGTCTGAATATCACGCTGAAGTCAGAAAGGTGCGAGATGCT 950
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIlyValArgAspSer 240
|
|
|
Db 951 TACCCAGACAGATGATGACACCATGTTCTGGCCGGTGACAAAGCAAGTAGAGCTCC 1010
QY 241 CysGlnGlyAspSerGlyGlyProValIlyCysAsnGlySerLeuGlnIlyLeuValSer 260
|
|
|
Db 1011 TGCACAGGTGATCTGGGGGGCTGTGTCTGCTGCAATGGCTCCCTGCAGGGACTCGTCTCC 1070
QY 261 TTPGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysIlyS 280
|
|
|
Db 1071 TGGGAGATTACCTTGTGCTCCGGCCCAACAGACCGGGTGTCTACGAACTCTGCAG 1130
QY 281 PheThrIlySTrpIleGlnIlyuThrIleGlnIlyAsnSer 293
|
|
|
Db 1131 TTCACCAAGTGATCCAGGAACCATCCAGGCCCACTCC 1169
RESULT 10
BD005362 1499 bp DNA linear PART 31-JAN-2002
LOCUS BD005362
DEFINITION Protease.
ACCESSION BD005362
VERSION BD005362.1 GI:18633733
KEYWORDS JP 2001501837-A/1.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1499)
AUTHORS Kitado,H., Yoshikawa,A. and Zaiki,T.
TITLE Protease
JOURNAL Patent: JP 2001501837-A 1 13-FEB-2001;
COMMENT THE PROCTER & GAMBLE CO
OS Human keratinocyte
PN JP 2001501837-A/1
PD 13-FEB-2001
PR 03-OCT-1997 JP 1999521662

PI HARUO KITADO, AKIKAZU YOSHIKAWA, TOMOKO ZAIKI
PC C12N15/57, C12N9/64, A61K38/48, A61K7/48, C11D3/386, C07K16/40 CC

FEATURES
source
1. .1499
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 1,42-120 Length: 1499
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-936-271B-14 (1-293) x BD005362 (1-1499)

QY 1 MetAlaThraIaArgProTpmETrpValLeuCyAlaLeuIleThraIaLeu 20
DB 291 ATGGCTACAGCAAGACCCCTGGATGTGGTCTGTCTGTGATCACAGCCCTGCTT 350

QY 21 LeuGlyValThrSlnIleValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 351 CTGGGGGTCAAGAGCATGTTCTGCCAACATGATGTTCTCTGTGACCACTCCCTCTAAC 410

QY 41 ThivAlProSerGlySerAsnGlnAspLeuGlyAlaGlyValAspAlaArgSer 60
DB 411 ACCGTCCTCTGGAGCAACAGAGCTGGAGCTGGGCGGAGAGAGCGCCGCTGC 470

QY 61 AsnAspSerSerAspGlyIleLeuAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 471 GATGACAGCAGAGCGGCATCATCAAGATGCCACTGCAATGACACCCAGCCGTGG 530

QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrcyGlyAlaValLeuValHisPro 100
DB 531 CAGGCGCGCTGTGCTAAGGCCCAACCACTCTACGCGGGCGGTGGTGGTCACTCA 590

QY 101 GlnTrpLeuLeuThraIaAlaHisCyAspGlyValValPheArgValaArgLeuGlyHis 120
DB 591 CAGTGGCTGCTCAAGCGCGCCCACTGCAGAGAAAGTTTTCAGAGTCCGCTCGGCAC 650

QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValValSerTlle 140
DB 651 TACTCCCTGTCCACAGTTTATGATCTGGGAGAGATGTTCCAGGGGGTCAAAATTCATC 710

QY 141 ProHisProGlyTyrSerHisAspProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
DB 711 CCCCACTCTGGCTACTCCACCTGGCCACTCTAACACCTCATGCTCATCAACTAAC 770

QY 161 ArgArgIleArgProThrIlyAspValaArgProIleAsnValSerSerHisCyAspProSer 180
DB 771 AGAAGAAATTCGTCCCACTTAAAGATGCAGACCCATCAACGCTCTCTCATTTGCCCTCT 830

QY 181 AlaGlyThrIlySerCyAlaLeuValSerGlyTTrpGlyThrThrIlySerProGlnValHisPhe 200
DB 831 GCTGGGCAAGAGCTGTGGTGTCTGGGAGCAACCAAGACCCCAAGAGTGCATTC 890

QY 201 ProIlyValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnIlyAspCyGluAspAla 220
DB 891 CTTAAGGTCTCTCCAGTCTTGAATATCAAGCTGTGAAGTCAAGAAAGTGGAGAGTCT 950

QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIlyAspIlyAspAspSer 240
DB 951 TACCCGAGACAGATGATGACACCATGTTCTGCGCCGAGTGAACAAAGCAGTGAAGATCTCC 1010

QY 241 CyGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 1011 TGCCAGGGGATTTCTGGGGGGCTGTGTGCTGCAATGGCTCCCTGCAGGAGCTCGTCC 1070

QY 261 TrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyGly 280
DB 1071 TGGGAGAGATTACCTTGTGCTGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTCAAG 1130

QY 281 PheThrIleTrpIleGlnIlyThrIleGlnAlaAsnSer 293
DB 1131 TTCACCAAGTGTATCCAGAAACCATTCAGGCCACTCC 1169

RESULT 11
BD107865
LOCUS BD107865 1516 bp DNA linear PAT 18-SEP-2002
DEFINITION 36 human secreted proteins.
ACCESSION BD107865
VERSION BD107865.1 GI:23202683
KEYWORDS JP 2002500035-A/36.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1516)
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,
Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.
TITLE 36 human secreted proteins
JOURNAL Patent: JP 2002500035-A 36 08-JAN-2002;
HUMAN GENOME SCIENCES INC
COMMENT OS Homo sapiens (human)
PN JP 2002500035-A/36
PD 08-JAN-2002
PR 06-JAN-1998 JP 2000527554
PR 07-JAN-1998 US 60/070657,07-JAN-1998 US 60/070658 PR
07-JAN-1998 US 60/070692,07-JAN-1998 US 60/070704 PI STEVEN
M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN
NI, LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K39/395, A61K48/00,
PC A61P5/00, A61P7/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC
A61P35/02,
PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC
C12N15/00,
PC A61K37/02, C12N5/00
CC n equals a, c, g, or c
FH Key Location/Qualifiers
FT source 1. .1516
FT location/Qualifiers
FT /organism="Homo sapiens"
FT /mol_type="genomic DNA"
FT /db_xref="taxon:9606"

FEATURES
source
1. .1516
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 1,42-120 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-936-271B-14 (1-293) x BD107865 (1-1516)

QY 1 MetAlaThraIaArgProTpmETrpValLeuCyAlaLeuIleThraIaLeu 20
DB 254 ATGGCTACAGCAAGACCCCTGGATGTGGTCTGTCTGTGATCACAGCCCTGCTT 313

QY 21 LeuGlyValThrSlnIleValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 314 CTGGGGGTCAAGAGCATGTTCTGCCAACATGATGTTCTCTGTGACCACTCCCTCTAAC 373

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValAspAlaArgSer 60
 Db 374 ACCGTCCTCTCGGAGCAACAGACCTCGGAGCTGGGGCCGGGAAGCCCGCTG 433
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
 Db 434 GATGACAGACAGACCGGCATCATCATGATCGACATGATGACATCCAGCCGTGG 493
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyCysGlyValAlaValLeuValHisPro 100
 Db 494 CAGGCGCGCTGTGCTAAAGCCCAACAGCTTACTGCGGGCGGTGGTGGCATCCA 553
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysAspGlyAlaValPheArgValArgLeuGlyHis 120
 Db 554 CAGTGCCTGCTACCGCGCCGCACTGACAGAAAGATTTCAGATCCGCTCGGCCAC 613
 QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnMetPheGlnGlyValTyrSerIle 140
 Db 614 TACTCCCTGCTACCACTTTATGATCTGGGCGACAGATGTTCCAGGGGGTCAATCCATC 673
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLysLeuAsn 160
 Db 674 CCCACCCCTGGCTACTCCACCTGGGCACTTAAACGACCTCATGCTCAATCAAACTGAAC 733
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
 Db 734 AGAAGAAATTCGTCCTCAATGAAGATGACAGACCATCACTCTCTCATTTCTCTCT 793
 QY 181 AlaGlyThrLysCysLeuValSerGlyTyrTrpGlyThrThrLysSerProGlnValHisPhe 200
 Db 794 GCTGGGACAAAGCTTGGTGTCTGGCTGGGAGCAACCAAGAGCCCAAGTGCACCTTC 853
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlyAspAla 220
 Db 854 CTTAAGCTCTCCAGTCTGATATCAGGAGTCTAAGTCAAGAAAGTGCAGGATGCT 913
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
 Db 914 TACCCGAGACAGATATGACACCATGTTCTGCGCGGTGACAAACAGATGAGACTCC 973
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
 Db 974 TCCAGAGGTGATTCCTGGGGGCTGCTGCTGCAAGGCTCCCTGACAGGACTCGATCC 1033
 QY 261 TrpGlyAspTrpProCysAlaIleArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
 Db 1034 TGGGGGATTAACCTTGTGCGGCCCAACAGACCGGGTGTCTACGAACTCTGCAAG 1093
 QY 281 PheThrLysTrpIleGlnGlnLysIleGlnAlaAsnSer 293
 Db 1094 TTCACCAAGTGAATCCAGGAACCATCCAGGCCAATCTCC 1132
 RESULT 12
 LOCUS BC008036 1527 bp mRNA linear PRI 29-JUN-2004
 DEFINITION Homo sapiens kallikrein 5, mRNA (cDNA clone MGC:1107
 IMAGE:2989806), complete cds.
 ACCESSION BC008036
 VERSION BC008036.2 GI:33873788
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1527)
 AUTHORS Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schaller, G.D.,
 Altschul, S.F., Zeeberg, B., Buecaw, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, J., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, J.,
 Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usslin, T.B., Toshiyuki, S.,

Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullan, S.J., Bogak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmitz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E.,
 Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 1527)
 STRAUSBERG, R.
 Direct Submission
 Submitted (21-MAY-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:14165477.
 CONTACT: MGC help desk
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadn@systemsbio.org
 Anup Madan, Jessica Fahy, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.linl.gov>
 Series: IRAL Plate: 3 Row: m Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 22208993.
 FEATURES
 source
 1..1527
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:1107 IMAGE:2989806"
 /issue_type="Ovary, adenocarcinoma"
 /clone_1fb="NIH MGC_9"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1..1527
 /gene="KLK5"
 /note="synonym: SCTB, KLK-L2, KLKL2"
 /db_xref="locusID:25818"
 /db_xref="MIM:605643"
 285..1166
 /gene="KLK5"
 /codon_start=1
 /product="kallikrein 5, preproprotein"
 /protein_id="AAH08036.1"
 /db_xref="GI:14165478"
 /db_xref="locusID:25818"
 /db_xref="MIM:605643"
 /translation="MATRRPFRMYTLCALTRLLIGTVEHYLANNDVSCDHSNTVPS
 GSNDDLAGAGSDADSDSSSTIINGSDCHDTQWQALILRPQLTCAVALVHPQW
 LITAAHCKRKYRVALGHVSLSPYVSGQWFOGVKSIPHPVSHRGSNDLMLTKLN
 RIRPRTKVRPIVNSHCPASGATKLVSGWGTTSPOVHPFVQLCLNISVLSQKRC
 DAYPQIDDTDFECADGRKGRDSCQSDSGPVVCSNLSGLVSWGVPARFRRPEVYT
 NLCKTKNIQETIQANS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.43e-120 Length: 1527

Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-936-271B-14 (1-293) x BC008036 (1-1527)

QY 1 MetA1aThrAlaArgProTpmETpValLeuCYaAlaLeu1eThrAlaLeu 20
DB 285 ATGGCTACAGAACACCCCTGATGTGGTGTCTGTCTGATCAACGCTTCTT 344
QY 21 LeuG1yValThrgLuh1sValLeuAlaAsnAAspValSerCYaAspHisProSerAsn 40
DB 345 CTGGGGGTCAAGAGCATTTCTGCGCAATATGTTCTCTGACCACTTCTTAC 404
QY 41 ThrValProSerGlySerAengInAAspLeuG1yAlaG1yAlaAspAlaArgSer 60
DB 405 ACCGTCCTCTGGAGAGCAACAGACCTGGAGCTGGGGCGGGAAGACGGCGTGC 464
QY 61 AspAAspSerSerAArg1le1leAeng1ySerAAspMetHisThgInProTpm 80
DB 465 GATACAGACAGACCCGATCATCATGATCGACATGCGATATGACACCCAGCCGTGG 524
QY 81 Glh1a1a1aLeuLeuLeuAArgProAengInLeuTyrCYaG1yAlaValLeuValHisPro 100
DB 525 CAGCCCGCTGTGCTTACAGCCCAACAGCTTACCTGCGGGCGGTGTGTCATCCA 584
QY 101 GlnTpmLeuLeuThrAlaAlaHisCYaArg1yAlaValPheArgValArgLeuG1yHis 120
DB 585 CAGTGGCTGTCAAGCGCGCCCACTGACAGAAAGTTTTCAGAGTCCGTCTGGCCAC 644
QY 121 TySerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValHisSer1le 140
DB 645 TACTCCCTGTACACAGTTTATGAAATCTGGGACAGATGTTCCAGGGGTCAAAATCCATTC 704
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAAspLeuMetLeu1leValLeuAsn 160
DB 705 CCCCACCTGTGCTTCTCCACCTGGCCACTGCTTACAGCTCATGCTCATTAACCTGAC 764
QY 161 ArgArg1leArgProThr1yAspValArgPro1leAsnValSerSerHisCYaProSer 180
DB 765 AGAAGATTCCTCCCACTTAAGATGACAGCCATCAACGTCTCTCATTTGCCCTCT 824
QY 181 AlaGlyThr1yScyValLeuValSerGlyTpmGlyThrThr1ySerProGlnValHisPhe 200
DB 825 GCTGGGACAAAGTCTGTGTCTGTGGTGGGACCAACAGAGCCCCCAAGTGCACATTTC 884
QY 201 ProValValLeuGlnCYaLeuAsn1leSerValLeuSerGlnValArgCYaGluAspAla 220
DB 885 CTTAAGGCTCTCCAGTGTCTTAATATACAGCTGTCTAAGTCAAGAAAGGTGCGAGATGCT 944
QY 221 TyrProArgGln1leAspAAspThrMetPheCYaAlaG1yAsp1yAlaG1yArgAAspSer 240
DB 945 TACCGAGACAGATAGATGACACCATGTTCTGGCGGCTGACAAAGCAGATAGAGACTCC 1004
QY 241 CyseGlnGlyAAspSerGlyGlyProValValCYaAeng1ySerLeuGlnGlyValHisSer 260
DB 1005 TGCCAGGAGTGAATCTGGGGGGCTGTGTGTCTGCAATGCTCTCCGACGGGACTGTGCTCC 1064
QY 261 TrpGlyAAspTyrProCYaAlaArgProAAspArgProGlyValTyrThrAAspLeuCYaHis 280
DB 1065 TGGGAGAGTTTACCTTGTGCGCCGACCAAGACCGGGGTGTCTACAGAACTCTGCAAG 1124
QY 281 PheThr1ySTp1leGlnGlnThr1leGln1a1aAsnSer 293
DB 1125 TTCACCAAGTGCATCAGAAACATTCAGGCGCAACTCC 1163

RESULT 13
AR525280 AR525280 1570 bp DNA linear PAT 20-DEC-2002
LOCUS AR525280
SEQUENCE 308 from patent US 6478825.
ACCESSION AR525280

VERSION AR525280.1 GI:27300488
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1570)
AUTHORS Winterbottom,J.M., Shimp,L., Boyce,T.M. and Kaes,D.
TITLE Implant, method of making same and use of the implant for the
JOURNAL treatment of bone defects
FEATURES Patent: US 6478825-A 308 12-NOV-2002;
source Location/Qualifiers
1..1570
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,48e-120 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-09-936-271B-14 (1-293) x AR525280 (1-1570)

QY 1 MetA1aThrAlaArgProTpmETpValLeuCYaAlaLeu1eThrAlaLeu 20
DB 354 ATGGCTACAGAACACCCCTGATGTGGTGTCTGTCTGATCAACGCTTCTT 413
QY 21 LeuG1yValThrgLuh1sValLeuAlaAsnAAspValSerCYaAspHisProSerAsn 40
DB 414 CTGGGGGTCAAGAGCATTTCTGCGCAATATGTTCTCTGTGACCACTTCTTAC 473
QY 41 ThrValProSerGlySerAengInAAspLeuG1yAlaG1yAlaAspAlaArgSer 60
DB 474 ACCGTGCTCTGGAGACACAGACCTGGAGCTGGGGCGGGAAGACGGCGTGC 533
QY 61 AspAAspSerSerAArg1le1leAeng1ySerAAspMetHisThgInProTpm 80
DB 534 GATGACAGACAGCCGACATCATCATGATCGATGTCACACCCAGCCGTGG 593
QY 81 Glh1a1a1aLeuLeuLeuAArgProAengInLeuTyrCYaG1yAlaValLeuValHisPro 100
DB 594 CAGCCCGCTGTGCTTACAGCCCAACAGCTTACCTGCGGGCGGTGTGTCATCCA 653
QY 101 GlnTpmLeuLeuThrAlaAlaHisCYaArg1yAlaValPheArgValArgLeuG1yHis 120
DB 654 CAGTGGCTGTCAAGCGCGCCCACTGACAGAAAGTTTTCAGAGTCCGTCTGGCCAC 713
QY 121 TySerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValHisSer1le 140
DB 714 TACTCCCTGTACACAGTTTATGAAATCTGGGACAGATGTTCCAGGGGTCAAAATCCATTC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAAspLeuMetLeu1leValLeuAsn 160
DB 774 CCCCACCTGTGCTTCTCCACCTGGCCACTTACAGCACTTCACTCATCAAACTGAC 833
QY 161 ArgArg1leArgProThr1yAspValArgPro1leAsnValSerSerHisCYaProSer 180
DB 834 AGAAGATTCCTCCCACTTAAGATGACAGCCATCAACGTCTCTCATTTGCCCTCT 893
QY 181 AlaGlyThr1yScyValLeuValSerGlyTpmGlyThrThr1ySerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGTCTGTGTGTGGTGGGACCAACAAAGCCCCCAAGTGCATTTC 953
QY 201 ProValValLeuGlnCYaLeuAsn1leSerValLeuSerGlnValArgCYaGluAspAla 220
DB 954 CTTAAGGCTCTCCAGTGTCTTAATATACAGCTGTCTAAGTCAAGAAAGTTCAGAGATGCT 1013
QY 221 TyrProArgGln1leAspAAspThrMetPheCYaAlaG1yAsp1yAlaG1yArgAAspSer 240
DB 1014 TACCGAGACAGATAGATGACACCATGTTCTGGCGCGGTGACAAAGCAGTAAAGACTCC 1073

QY 241 CyGInGIYAspSerGIYgIyProValIaIaCyAsnGIYSerLeuInGIYLeuValIser 260
|
|
|
Db 1074 TGGCAGGGATGATTCGGGGGGGCGTGTGTCTGCATAGGCTCCTCGCAGGAGACTCGTGTCC 1133
|
|
|
QY 261 TTPGIYAspTYrProCYaIaArpProAsnArpProGIYValIYrThraenLeuCYaIys 280
|
|
|
Db 1134 TGGGAGATTACCTGTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1193
|
|
|
QY 281 PheThrIYerTrpIleGIInGIYrIleGIInIaAsnSer 293
|
|
|
Db 1194 TTCACCAAGTGCATCCAGAAACCATCCAGGCCAATCTCC 1232
|
|
|
RESULT 14
AR528679 1570 bp DNA linear PAT 08-OCT-2004
LOCUS Sequence 455 from patent US 6725730.
DEFINITION AR528679
VERSION AR528679.1 GI:53916757
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 1570)
TITLE Bollinger, C.L. Jr.
JOURNAL Crane test weight assembly and method
FEATURES Patent: US 6725730-A 455 27-APR-2004;
Location/Qualifiers
1..1570
Source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 1,48e-120 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-936-271B-14 (1-293) x AR528679 (1-1570)

QY 1 MetalAlThrAlaArpProTrpMetTrpValIeucYsaIaLeuIleThraIaLeu 20
|
|
|
Db 354 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTGTCTGTCTGTATCAAGCTTGCTT 413
|
|
|
QY 21 LeuGIYValIThrGIuIhIValIeuaIaAsnAsnAspValSerCYaAspHIsProSerAsn 40
|
|
|
Db 414 CTGGGGGTCAACAGCATGTTCTGCGCAACATGATGTTCTGTGACCAACCCCTTAAC 473
|
|
|
QY 41 ThrValProSerGIYSerAsnGIYnaAspLeuGIYValIaGIYValIaGIYValIaAspSer 60
|
|
|
Db 474 ACCGTGCTCTGTGAGCAACAGACCTGTGAGCTGGGGCGGGGAGACGGCGGTG 533
|
|
|
QY 61 AspAspSerSerSerArgIleIleAsnGIYSerAspCYaAspMetHIsThrGIInProTrp 80
|
|
|
Db 534 GATGACAGCAGCAGCCGATCATCATGATCGACATGATGACACCCAGCCGTGG 593
|
|
|
QY 81 GluAlaIaLeuLeuLeuValProAsnGIYLeuTYrCYaGIYValIaValIeValHIsPro 100
|
|
|
Db 594 CAGGCGCGCTGTGTGCTAAGGCCAACCACTCACTGCGGGCGGGGTGTGGTGCATCCA 653
|
|
|
QY 101 GluTrpLeuLeuThrAlaAlaHIsCYaArGIYValIaPheArgValIaArgLeuGIYHIs 120
|
|
|
Db 654 CAGTGGCTGTCTACAGCGCGCCACTGCAGAGAAAGATTTCAGAGTCCGTCCGCGCAC 713
|
|
|
QY 121 TyrSerLeuSerProValIYrGIuSerGIYGIInIaMerPheGIInGIYValIYsSerIle 140
|
|
|
Db 714 TACTCCTGTACACAGTTTATGATCTGGGACAGCATGTTCCAGGGGGGTCAATTCATC 773
|
|
|
QY 141 ProHIsProGIYTYrSerHIsProGIYHIsSerAsnAspLeuMetLeuIleYsLeuAsn 160
|
|
|

Db 774 CCCCACTGTGCTACTCCCACTGGGCCACTTAACGACCTCATGCTCAACTGAAC 833
|
|
|
QY 161 ArgArgIleArGIProThrIYsaAspValaArgProIleAsnValSerSerHIsCYaProSer 180
|
|
|
Db 834 AGAAGATTGCTCCACTTAAGATGTCAGACCCATCAAGTCTCTCTCATTTGTCTCT 893
|
|
|
QY 181 AlagIYThrIYCySLeuValSerGIYTYrGIYrThrThIYSerProGIInValHIsPhe 200
|
|
|
Db 894 GCTGGGACAAAGTGTGTGTGTGTGCTGGCTGGGGGACAAACAGAGGCCCAAGTGCATTC 953
|
|
|
QY 201 ProIYsValIeGIInCYaLeuAsnIleSerValIeSerGIInIYArGIYsGIuAspAla 220
|
|
|
Db 954 CCTAAGTCTCTCAGGCTTGAATATACAGCGGTCTAAGTCAAGAAAGGTCCAGAGTGT 1013
|
|
|
QY 221 TYrProArgGIInIleAspAspThrMetPheCYaAlaGIYAspIYsaIaIYrAspSer 240
|
|
|
Db 1014 TACCCGAGACAGATATATACACCATGTCTGTGGCGGAGACAAACAGATGAGACTCC 1073
|
|
|
QY 241 CyGInGIYAspSerGIYgIyProValIaIaCyAsnGIYSerLeuInGIYLeuValIser 260
|
|
|
Db 1074 TGGCAGGGATGATTCGGGGGGGCGTGTGTCTGCATAGGCTCCTCGCAGGAGCTCGTGTCC 1133
|
|
|
QY 261 TTPGIYAspTYrProCYaIaArpProAsnArpProGIYValIYrThraenLeuCYaIys 280
|
|
|
Db 1134 TGGGAGATTACCTGTGTGCGCGGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1193
|
|
|
QY 281 PheThrIYerTrpIleGIInGIYrIleGIInIaAsnSer 293
|
|
|
Db 1194 TTCACCAAGTGCATCCAGAAACCATCCAGGCCAATCTCC 1232
|
|
|
RESULT 15
AX080829 1570 bp DNA linear PAT 27-FEB-2001
LOCUS AX080829
DEFINITION Sequence 75 from Patent WO0109327.
ACCESSION AX080829
VERSION AX080829.1 GI:13169796
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Ashkenazi, A., J., Baker, K. P., Goddard, A., Godowski, P. J., Gurney, A. L.,
Klajavin, I. J., Lafleur, M., Mark, M. R., Marsters, S. A., Pitti, R. M.,
Matanabe, C. K. and Wood, W. I.
TITLE Method of preventing the injury or death of retinal cells and
JOURNAL treating ocular diseases
FEATURES Patent: WO 0109327-A 75 08-FEB-2001;
Genentech, Inc. (US)
SOURCE Location/Qualifiers
1..1570
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 1,48e-120 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-936-271B-14 (1-293) x AX080829 (1-1570)

QY 1 MetalAlThrAlaArpProTrpMetTrpValIeucYsaIaLeuIleThraIaLeu 20
|
|
|
Db 354 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTGTCTGTCTGTATCAAGCTTGCTT 413
|
|
|
QY 21 LeuGIYValIThrGIuIhIValIeuaIaAsnAsnAspValSerCYaAspHIsProSerAsn 40
|
|
|
Db 414 CTGGGGGTCAACAGCATGTTCTGCGCAACATGATGTTCTGTGACCAACCCCTTAAC 473
|
|
|

```
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValIleGlyValAspAlaArgSer 60
DB 474 ACCGTGCTCTTGGAGACAAACGAGACTGGAGCTGGAGCCGGAGAGAGCCCGGTTCG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 534 GATGACAGCAGACGCCCATCATCATGATCGACTGGCATATGCAACCCAGCCGTGG 593
QY 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValIleValLeuValHisPro 100
DB 594 CAGGCCCGCTGTGTCTAAGGCCCAACAGCTTACTGCGGCGCGGTGTGTGCATCCA 653
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgIleValIleValIlePheArgValArgLeuGlyHis 120
DB 654 CAGTGGCTGCTCAGCGCGCCGCTGCAAGAGAAAGTTTTCAGAGTCCGTTCTGGCCAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 714 TACTCCCTGTACCAAGTTTATGATCTGGGACAGAGATGTTCCAGGGGGTCAATCCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
DB 774 CCCCACTGCTGCTACTCCACCCCTGGCCACTCTAACGACCTCATGCTCAACAACTGAAC 833
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 834 AGAAGAATGTGCTCCACTTAAGATGACAGCCCAACAGCTCTCTCATGTGCTCTCT 893
QY 181 AlIGlyThrIleCysLeuValIleSerGlyTyrGlyTyrThrIleCysSerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGTCTGTGTCTGTCTGGGCGGACAAACAAAGCCCCCAAGTCACTTC 953
QY 201 ProIleValIleGlnCysLeuAsnIleSerValIleSerGlnIleValArgCysGluAspAla 220
DB 954 CCTAAGGTCTCTCCAGTCTTGAATATACGCGCTGAAGTCAAGAAAGGTGCGAGATGCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValIleArgAspSer 240
DB 1014 TACCGAGACAGATGATGACACCATGTCTGCGCGCGGTGACAAAGCAGATGAGACTCC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TGGCAGGGTGAATCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAAGGACTGTCTCC 1133
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValIleTyrThrAsnLeuCysLeu 280
DB 1134 TGGGAGATTAACCTTGTGCCCCGCCCAACAGACCGGGGTGTCTACAGAACTCTGCAAG 1193
QY 281 PheThrIleTyrTrpIleGlnGlnIleThrIleGlnAlaAsnSer 293
DB 1194 TTCAACCAAGTGATCCAGAAACCATCCAGGCCAACTCC 1232
```

Search completed: February 26, 2005, 21:24:06
Job time : 4922 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 11:08:50 ; Search time 609 Seconds

(without alignments)
2848.086 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608
Sequence: 1 MATARPPPMVLCALITALL.....VTNLCCKFTKTIQETIQANS 293

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%

Listing first 45 summaries

Command line parameters:

-MODEL=Frame+g2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.spool/US09936271/runat.23022005.130126.27251/app_query.fasta_1.455
-DB=Geneseq.16Dec04 -OPMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09936271 @CGN_1_1_470 @runat.23022005.130126.27251 -ACPU=6 -ICPU=3
-NO_MMAP -JARGOVERLY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	882	2	AAZ23318 Human PDS
2	1608	100.0	1216	8	ADA39888 Human sec
3	1608	100.0	1216	10	ADA56078 Gene enco
4	1608	100.0	1260	10	ADB80489 Ovarian c
5	1608	100.0	1260	11	ADN39197 Cancer/an

6	1608	100.0	1260	13	ADR25550	Adi25550 Breast ca
7	1608	100.0	1302	2	AA57999	Aax57999 Human BS2
8	1608	100.0	1302	2	AA57989	Aax57989 Human BS2
9	1608	100.0	1381	2	AA206259	Aaz06259 Human sec
10	1608	100.0	1381	10	ADA40370	Ada40370 Human sec
11	1608	100.0	1381	10	ADA56535	Ada56535 Gene enco
12	1608	100.0	1387	13	ADR72632	Adi72632 Human ren
13	1608	100.0	1387	13	ADR72874	Adi72874 Human ova
14	1608	100.0	1490	8	ADA40367	Ada40367 Human sec
15	1608	100.0	1490	10	ADA56532	Ada56532 Gene enco
16	1608	100.0	1499	2	AA560578	Aax60578 Human ker
17	1608	100.0	1516	2	AA206245	Aaz06245 Human sec
18	1608	100.0	1516	8	ADA40369	Ada40369 Human sec
19	1608	100.0	1516	10	ADA56534	Ada56534 Gene enco
20	1608	100.0	1539	2	AA23317	Aaz23317 Human PDS
21	1608	100.0	1570	3	AA265070	Aaz65070 Membrane-
22	1608	100.0	1570	4	AA521471	Aa521471 Human CDN
23	1608	100.0	1570	5	AA44216	Aa44216 Human PRO
24	1608	100.0	1570	6	ABK28605	Abk28605 Human DNA
25	1608	100.0	1570	8	ACA4363	Aca4363 Novel hum
26	1608	100.0	1570	8	ACA03830	Aca03830 CDNA enco
27	1608	100.0	1570	8	ABX89368	Abx89368 DNA encod
28	1608	100.0	1570	8	ABX80822	Abx80822 Human sec
29	1608	100.0	1570	8	ACD44331	Ac44331 CDNA enco
30	1608	100.0	1570	8	ACD42022	Ac42022 Human sec
31	1608	100.0	1570	8	ABX79502	Abx79502 Human sec
32	1608	100.0	1570	8	ACA93523	Aca93523 Novel hum
33	1608	100.0	1570	8	ABX81205	Abx81205 Human sec
34	1608	100.0	1570	8	ACA04251	Aca04251 Human CDN
35	1608	100.0	1570	8	ACA93021	Aca93021 Novel hum
36	1608	100.0	1570	8	ABX17105	Abx17105 Human PRO
37	1608	100.0	1570	9	ACA57960	Aca57960 Novel hum
38	1608	100.0	1570	9	ACA88409	Aca88409 Human sec
39	1608	100.0	1570	9	ACD81916	Ac481916 CDNA enco
40	1608	100.0	1570	9	ADA45974	Ada45974 Novel hum
41	1608	100.0	1570	9	ADA76405	Ada76405 Human PRO
42	1608	100.0	1570	9	ADA19055	Ada19055 Human PRO
43	1608	100.0	1570	9	ADA51678	Ada51678 Homo sapl
44	1608	100.0	1570	9	ADB19463	Adb19463 Novel hum
45	1608	100.0	1570	9	ADB28004	Adb28004 CDNA enco

ALIGNMENTS

RESULT 1	
AAZ23318	
ID	AAZ23318 standard; DNA; 882 BP.
XX	
AC	AAZ23318;
XX	
DT	03-DEC-1999 (first entry)
XX	
DE	Human PDSF-1 DNA coding region.
XX	
KM	PDSF-1; Tango 114; prostate-derived serine protease; fusion protein;
KM	treatment; proliferative disorder; prostate cancer; screening;
KM	cellular proliferation; growth factor; growth factor biosynthesis;
KM	cellular differentiation; growth factor binding protein; metastasis;
KM	cellular differentiation; prostate development; detection; forensic;
XX	serine protease; human; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO946391-A2.
XX	
PD	16-SEP-1999.
XX	
PF	12-MAR-1999; 99WO-US005416.
XX	
PR	12-MAR-1998; 98US-00041400.
XX	
PA	(MIL-) MILLENNIUM BIOTHERAPEUTICS INC.

PI Holtzman DA;
 XX WPI; 1999-561678/47.
 DR P-PSDB; AAY30524.
 XX
 PT Novel prostate derived serine protease polynucleotides and polypeptides
 XX used to modulate cellular processes.
 XX
 PS Claim 1a; Fig 1A; 94pp; English.
 XX
 CC This invention describes a novel human prostate-derived serine protease
 CC (PDSp) polynucleotide (also known as Tango 114). The prostate derived
 CC serine protease (PDSp) polypeptide is used to modulate a variety of
 CC cellular processes. It can be used to produce fusion proteins. PDSp is used
 CC to treat proliferative disorders, e.g. prostate cancer. The protein may
 CC also be used to produce antibodies, and to identify antagonists and
 CC agonists. The PDSp polynucleotides, polypeptides, homologs and antibodies
 CC can be used in screening assays; predictive medicine; and methods of
 CC treatment. PDSp cleaves growth factors, and can be used for the
 CC modulation of growth factor biosynthesis; generation of active peptides;
 CC regulation of cellular proliferation; degradation of growth factor
 CC binding proteins; regulation of cellular differentiation; regulation of
 CC metastasis; and regulation of prostate development. The PDSp
 CC polynucleotides can be used to express the protein; to detect PDSp mRNA;
 CC to detect genetic alterations in the PDSp gene; in forensic biology; and
 CC as a source of primers and probes. As serine protease have important
 CC roles in cellular processes, there exists a need for identifying novel
 CC serine protease such as the prostate derived serine protease (PDSp) of
 CC the invention. This sequence encodes the human PDSp-1 described in the
 CC invention
 XX
 SQ Sequence 882 BP; 193 A; 272 C; 237 G; 180 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 7.44e-136 Length: 882
 Score: 1608.00 Matches: 293
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0
 US-09-936-271B-14 (1-293) x AA223318 (1-882)
 QY 1 MetAlaThrAlaArgProTyrMetTyrValLeuCyAlaLeuIleThrAlaLeu 20
 Db 1 ATGGCTACAGCAAGCCCTCGATGTGGTCTCTGTGATCAAGCTTGCTT 60
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
 Db 61 CTGGGGGTCAAGAGCATGTTCTGCGCAACATGATGTTCCGTGACCAACCTCTAAC 120
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGluAspAlaArgSer 60
 Db 121 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGAGAGCGCCGATG 180
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80
 Db 181 GATGACAGACGACGCGCATCATCATGATGATCCGATTCGATTCACACCCGCGTGG 240
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValValLeuValHisPro 100
 Db 241 CAGGCGCGCTGTGTGCTAAAGCCCAACAGCTCACTGCGGGGTGTGGTGCATCCA 300
 QY 101 GlnThrLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
 Db 301 CAGTGGCTGTCAAGCGCGCCACTGACGAGAAAGATTTCAGAGTCCCTCGGCCAC 360
 QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
 Db 361 TACTCCCTGTCAACAGTTTATGATCTGGGACACAGATGTTCCAGGGGGTCAATCATC 420
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160

Db 421 CCCACCTGGCTACTCCACCCCTGGCCACTGTAACGACCTCATGCTCAAACTGAAC 480
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCyProSer 180
 Db 481 AGAAGATTGCTCCCACTAAAGATGTCAAGCCCATCAAGCTCTCTCTATTGTCCTCT 540
 QY 181 AlaGlyThrLysCyLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
 Db 541 GCTGGGACAAAGTGTGTGTGCTGGCTGGGGACAAACCAAGAGCCCCAAGTGCCTTC 600
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgGlyAspAla 220
 Db 601 CTTAAGTCTCTCAAGCTTGAATATCAACCGGTCTAAGCAAAAAGGCGACGATGCT 660
 QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspLysValGlyArgAspSer 240
 Db 661 TACCCGAGCAATATGATATACCATGTTTCTGCGCGGTGACAAAGCATGTAGACTTC 720
 QY 241 CyGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
 Db 721 TGCAGGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTGACGGGACTCGTCTC 780
 QY 261 TrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyLys 280
 Db 781 TGGGGAGATTACCTTGTGCTCCGGCCCAACACACCGGGGTGTACACGAACCTTGCAAG 840
 QY 281 PheThrIleTyrIleGlnIleThrIleGlnAlaAsnSer 293
 Db 841 TTCACCAAGTGAATCCAGGAACATTCACGGCCAACTCC 879
 RESULT 2
 ADA39888
 ID ADA39888 standard; cDNA; 1216 BP.
 XX
 AC ADA39888;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein encoding cDNA.
 XX
 KW Human; secreted protein; cancer; hyperproliferative disorder;
 KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
 KW anaemia; allergic reaction; asthma; cardiovascular disorder;
 KW wound healing; cytoskeletal; immunosuppressive; neuroprotective;
 KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
 KW vulnerability; cardiact; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO2002102993-A2.
 XX
 PD 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002MO-US008123.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCT INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-175238/17.
 XX
 PT New human secreted proteins and nucleic acid molecules, useful for
 PT preventing a diagnostic or pharmaceutical composition for diagnosing,
 PT preventing or treating cancer or other hyperproliferative disorder,
 PT asthma, allergies or AIDS.
 XX
 PS Claim 9; SEQ ID NO 270; 3205bp; English.
 XX
 CC The invention relates to novel genes ADA39629-ADA40565 and proteins

CC ADA40566-ADA1501 for human secreted proteins, useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind to the polypeptide are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating cancer or other hyperproliferative disorder. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating cancer or other hyperproliferative disorders including neoplasms, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic anaemia), haematopoietic or haematological disorders (e.g. anaemia, thrombocytopenia), allergic reactions including asthma or eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory bowel disease or Crohn's disease), neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's disease), cardiovascular disorders (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial, fungal or viral infections including HIV/AIDS), or wound healing and disorders of epithelial cell proliferation. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1216 BP; 263 A; 388 C; 323 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,136-135	Length:	1216
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-936-271b-14 (1-293) x ADA39888 (1-1216)

QY 1 MetLathra1aAaGProProTmMetTrrYalLeuCySa1aLeu1leThra1aLeu1leu 20
DB 128 ATGCTACAGAAACCCCTGATGTGGTGTCTGTCTGATCAGCAGCTTCTT 187
QY 21 Leu1YalThrg1uH1eValLeu1aAaAaAaAaAaAaAaAaAaAaAaAaAaAaAa 40
DB 188 CTGGGGGTCAAGAGCATGTCTCGCAACATGATGTTCTCTGTGACCAACCCCTTAC 247
QY 41 Thra1ProserGysaAa 60
DB 248 ACCGTGCTCTGGAAGAACAGAGACCTGGAGCTGGGGGGAAGAGCCCGGTG 307
QY 61 Aa 80
DB 308 GATGACAGAGCAACCCCATCATCATGATCGACATGCAATGCAACCCAGCCGTG 367
QY 81 Gln1a1a1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e 100
DB 368 CAGCCCGCGCTGTGCTAAAGCCCAACAGCTTACGCGGGGGGTGTGTGTGATCA 427
QY 101 Gln1r1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e 120
DB 428 CAGTGGCTGCTCAAGCGCGCCACTGACAGAAAGATTTCAGAGTCCGCTGGCCAC 487
QY 121 TySer1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e 140
DB 488 TACTCCCTGTCACAGTTTATGAATCTGGGACAGATGTTCCAGGGGGGTCAATTCATC 547
QY 141 ProH1aProG1y 160
DB 548 CCCCAACCTGCTCTCCACCCCTGGGACACTTAAAGCACTCAAGCTCATCAAACTGAC 607
QY 161 ArgArg1e1aAaGProThr1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y1y 180

DB 608 AGAAGATTGCTCCCACTAAAGATGTCAGCCATCAAGCTCTCTCATTTGCCCTCT 667
QY 181 AlG1yThry1y 200
DB 668 GCTGGGCAAAAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 727
QY 201 Pro1y1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e1u1e 220
DB 728 CTTAAGTCTCTCAAGT 787
QY 221 TyPro1y 240
DB 788 TACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 847
QY 241 CyG1y 260
DB 848 TGGCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 907
QY 261 TrG1y 280
DB 908 TGGGAGATTAACCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 967
QY 281 PheTh1y 293
DB 968 TTCACCAAGTGAATCCAGGAACATCCAGGCCACTCC 1006

RESULT 3

ADA56078 ID ADA56078 standard; DNA; 1216 BP.

XX AC ADA56078;

XX XX 20-NOV-2003 (first entry)

XX XX Gene encoding human secreted protein #257.

XX XX immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
XX XX cytoprotective; cerebroprotective; neuroprotective; nootropic;
XX XX cardiovascular; antiarteriosclerotic; gene therapy;
XX XX human secreted protein; immune disorder; inflammation;
XX XX respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
XX XX inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
XX XX multiple sclerosis; ischemic brain injury; Parkinson's disease;
XX XX Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
XX XX triple helix formation; antisense gene therapy; forensic biology; ds;
XX XX gene.

XX OS Homo sapiens.

XX XX W02002102994-A2.

XX XX 27-DEC-2002.

XX XX 19-MAR-2002; 2002WC-US008278.

XX XX 21-MAR-2001; 2001US-0277340P.

XX XX 19-JUL-2001; 2001US-0306171P.

XX XX 13-NOV-2001; 2001US-0331287P.

XX XX (HUMA-) HUMAN GENOME SCI INC.

XX XX Rosen CA, Ruben SM;

XX XX WPI; 2003-167512/16.

XX XX P-PSDB; ADA56974.

XX XX New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.

PS Claim 21; SEQ ID NO 267; 1754bp; English.

XX The invention relates to 592 new human secreted polypeptides useful for
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
 CC conditions, respiratory disorders, cancers, CNS disorders, or
 CC neurodegenerative disorders, or polypeptides comprising an amino acid
 CC sequence at least 95% identical to the new sequences. The polypeptides,
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic
 CC acids encoding the polypeptides, agonists or antagonists that binds to
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
 CC compositions for diagnosing, treating or preventing an e.g. immune
 CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 CC polynucleotides are useful for chromosome identification, chromosome
 CC mapping, for controlling gene expression through triple helix formation
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals
 CC from minute biological samples, in forensic biology, and as hybridization
 CC probes. The polypeptides are useful for as molecular weight markers on
 CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 CC gels, to raise antibodies, for testing biological activities, and for
 CC treating or preventing neural disorders, immune system disorders,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds
 CC to a gene encoding one of the polypeptide of the invention. Note: The
 CC sequence data for this patent did form part of the printed specification,
 CC but was obtained in electronic format directly from WIPD at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1216 BP; 263 A; 388 C; 323 G; 242 T; 0 U; 0 Other;

Alignment Scores:

Score: 1.13e-135 Length: 1216
 Percent Similarity: 1608.00 Matches: 293
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 100.00% Mismatches: 0
 DB: 10 Gaps: 0

US-09-936-271B-14 (1-293) x ADA56078 (1-1216)

QY 1 MetAlaThrAlaArgProPrtMetTrrValLeuCyAlaLeuIleThrAlaLeu 20
 Db 128 ATGGCTACAGCAAGACCCCTGGATGGGTCTCTGTGCTGATCACAGCTTGCTT 187
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
 Db 188 CTGGGGGTCAACAGCATGTTCTGGCCAACAATGATGTTCCGTGACCAACCCCTCTAAC 247
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyValAlaGlyIleuAspAsp 60
 Db 248 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAAGCGCCGCTG 307
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrr 80
 Db 308 GATGACAGGACGACCGCATCATCATGATCGACATGATGACACCCAGCGCTGG 367
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrrCysGlyAlaValLeuValHisPro 100
 Db 368 CAGGCGCGCTGTGTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGTGTCATCA 427
 QY 101 GlnTrrLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
 Db 428 CAGTGGCTGCTACCGGCGCCCACTGAGAGAAAGATTTCAGAGTCCGTCTCGGCCAC 487
 QY 121 TyrSerLeuSerProValTrrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
 Db 488 TACTCCCTGTACCAAGTTATGAAATCTGGGACGACATGTTCCAGGGGGGTCAAAATCATC 547
 QY 141 ProHisProGlyTrrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160

Db 548 CCCACCTGTGCTACTCCACCCTGGCACTCACTAAAGACCTCATCTCAAACTGAAC 607
 QY 161 ArgArgIleAlaArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
 Db 608 AGAAGAAATTCCTCCCACTAAAGATGCAGACCATCAAGCTCTCTCATTTGCTCTCT 667
 QY 181 AlaGlyThrIleCysLeuValSerGlyTrrGlyThrThrIleSerProGlnValHisPhe 200
 Db 668 GCTGGACAAAGTCTGGTGTCTGGCTGGGGGACACCAAGACCCCAAGTGCATTC 727
 QY 201 ProIleValLeuGlnCysLeuAsnIleSerValIleSerGlnIleArgCysGluAspAla 220
 Db 728 CTTAAGTCTCTCAAGCTTGAATATCAACGCTGCTAAAGTCAAGAAAGGTCCAGATGCT 787
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspValAlaIleArgAspSer 240
 Db 788 TACCCAGACAGATGATGATACACCATGTTCTGCGCGGACAAAGCAGTGAAGCTCC 847
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
 Db 848 TGCAGGGGTGATTTGGGGGGCTGTGTGCTGCATGCTCCCTCGAGGAGCTCGTGTCC 907
 QY 261 TrrGlyAspTrrProCysAlaArgProAsnArgProGlyValIleTrrAsnLeuCysIle 280
 Db 908 TGGGGGATTAACCTTGTGCTCCGCGCCAAACAGACGGGGTGTACACGAACCTTGCAAG 967
 QY 281 PheThrIleTrrIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 293
 Db 968 TTCACCAAGTGAATCCAGGAAACCATCCAGGCCCAACTCC 1006

RESULT 4

ADB80489
 ID ADB80489 standard; DNA; 1260 BP.

AC ADB80489;

DT 04-DEC-2003 (first entry)

DE Ovarian cancer-associated transcript #27.

KW cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;

KW post-operative chemotherapy; radiation therapy; tumour prognosis;

KW pre-cancerous lesion detection; de; gene.

OS Homo sapiens.

PN MO2002102235-A2.

PD 27-DEC-2002.

PF 18-JUN-2002; 2002WO-US019297.

PR 18-JUN-2001; 2001US-0299234P.

PR 27-AUG-2001; 2001US-0315287P.

PR 05-SEP-2001; 2001US-0317544P.

PR 13-NOV-2001; 2001US-0350666P.

PR 12-APR-2002; 2002US-0372246P.

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Mack DH, Gish KC;

DR WPI; 2003-167431/16.

PT P-PSDB; ADB80490.

PT Detecting an ovarian cancer-associated transcript in a cell from a
 PT patient; comprises contacting a biological sample from the patient with a
 PT polynucleotide that hybridizes to an ovarian cancer gene.

PS Claim 10; Page 292; 332pp; English.

XX The invention relates to a method of detecting an ovarian cancer-associated transcript in a cell from a patient, by contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to any of one of 80 nucleic acid sequences given in the specification. The method is useful in diagnosing ovarian cancer and in identifying and using agents and/or targets that inhibit ovarian cancer. The nucleic acid molecule, polypeptide, and the antibody may also be used in detecting ovarian cancer, monitoring and early detection of relapse following treatment, monitoring response to therapy, selecting patients for post-operative chemotherapy or radiation therapy, in selecting mode of therapy, determining tumour prognosis, early detection of pre-cancerous lesions, and as vaccines. This sequence corresponds to one of the nucleic acids used for the detection method of the invention.

SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.18e-135	Length:	1260
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-936-271b-14 (1-293) x ADB80489 (1-1260)

QY	1	MeAlaThraLaArpProPOTrMetTrValLeuCYaAlaLeuIleThraAlaLeu	20
DB	43	ATGGCTACAGACAGACCCCTGTGATGTGGTCTGTGTCTGATCAGACGCTTCTT	102
QY	21	LeuGIYValThrgIuHISValLeuAlaAsnAsnAspValSerCYaAspHISProSerAsn	40
DB	103	CTGGGGGTCAAGAGCATGTTCTGCAACATATATTTCTGTGACCAACCCCTTAAAC	162
QY	41	ThrValProSerGIYSerAsnGlnAspLeuGIYValaGIYValaGIYValaAspAsp	60
DB	163	ACCGTGGCTGTGGAGCAACAGACCTGGAGCTGGGGGGGAGAGCGCGGGTGG	222
QY	61	AspAspSerSerSerArgIleIleAsnGIYSerAspCYaAspMetHISThrGlnProTrp	80
DB	223	GATGACAGCAGCAGCCCATCATCAATGATCCAGTCCGATATGCAACCCAGCCCTGG	282
QY	81	GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTrCYaGIYValaValaLeuValHISPro	100
DB	283	CAGGCGCGCTGTGGCTAAAGGCCCAACAGCTTACTGGGGGGGGTGTGTGCATCCA	342
QY	101	GltTrpLeuLeuThraAlaAlaHISCYaArgIYValaValaPheArgValaIleLeuGIYHIS	120
DB	343	CAGTGGGTGTCAAGCGCGCCCATGTGAGAGAAAGTTTTCAGAGTCCGTCTGGGCAC	402
QY	121	TyrSerLeuSerProValTYrGIYSerGIYGlnGlnMetPheGlnGIYValaIYSerIle	140
DB	403	TACTCCCTGTCAACAGTTTATGATCTGGGCAAGAGATTTCCAGGGGGGTCAATTCATC	462
QY	141	ProHISProGIYTYrSerHISProGIYHISerAsnAspLeuMetLeuIleYValaAsn	160
DB	463	CCCCACCCGTGTACTCCACCCCTGGCCACTCTAAACAGCTCATGCTCAACAACTGAAC	522
QY	161	ArgArgIleArgProTrpTrpAspValArgProIleAsnValISerSerHISCYeProSer	180
DB	523	AGAGAAATGTCTCCACTAAAGATGACAGACCCATCAACGCTCTCTCATGTGCCCTCT	582
QY	181	AlaGIYThryrCYaLeuValISerGIYTrpGIYThrThryrSerProGlnValHISPhe	200
DB	583	GCTGGGAGCAAGTCTGTGTGTGTGGTGGGGAGCAACAGAGCCCCCAAGTCACTTC	642
QY	201	ProIYValLeuGlnCYaLeuAsnIleSerValISerGlnIYValaValaCYaGIYAspAla	220
DB	643	CCTAAAGTCTCCAGTCTTGAATATCAAGGCTGCTAAGTCAAGAAAGTGGAGAGTCT	702

QY	221	TyrProArgGlnIleAspAspThrMetPheCYaAlaGIYAspIYValaIYArgAspSer	240
DB	703	TACCCAGACAGATATATATACACCATCTTCTGGCCGGGTGACAAACAGTTAAGACTCC	762
QY	241	CYeGlnGIYAspSerGIYGIYProValaValaCYaAsnGIYSerLeuGlnIYLeuValISer	260
DB	763	TGCCAGGGTGATCTGGGGGGGCTGTGTCTGCAATGGCTCCCTGCAGGAGACTCGTTC	822
QY	261	TrpGIYAspTYrProCYaAlaArgProAsnArgProGIYValaTYrThraMetLeuCYeIYs	280
DB	823	TGGGAGATTAACCTTGTGTCCGCCCAACAGCCGGGTGTCTAACAGAACTTGTCAAG	882
QY	281	PheThryrTrpIleGlnIuTrIleGlnAlaAsnSer	293
DB	883	TTCAACAGTGTATCCAGAAACATTCAGGCCAATCTCC	921

RESULT 5

ADN39197

ID ADN39197 standard; cDNA; 1260 BP.

AC ADN39197;

DT 17-JUN-2004 (first entry)

DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:515.

KW Human; differential expression; cancer; angiogenic disorder;

KW fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis;

KW inflammatory disease; autoimmune disease;

KW retinal neovascularization syndrome; scarring; uterine fibroid;

KW wound healing; contraindication; cytostatic; cardiact; immunomodulatory;

KW vulnerable; gene therapy; vaccine; gene; ss.

OS Homo sapiens.

PN WO2003042661-A2.

PD 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

XX 21-NOV-2001; 2001US-0332464P.

XX 29-NOV-2001; 2001US-0334393P.

XX 03-DEC-2001; 2001US-0340376P.

XX 08-JAN-2002; 2002US-0347211P.

XX 10-JAN-2002; 2002US-0347349P.

XX 08-FEB-2002; 2002US-035250P.

XX 13-FEB-2002; 2002US-0356714P.

XX 20-FEB-2002; 2002US-0359077P.

XX 29-MAR-2002; 2002US-036809P.

XX 04-APR-2002; 2002US-0370110P.

XX 12-APR-2002; 2002US-0372246P.

XX 05-JUN-2002; 2002US-0386514P.

XX 16-JUL-2002; 2002US-0396839P.

XX 22-JUL-2002; 2002US-039775P.

XX 09-SEP-2002; 2002US-0409450P.

PA (E0SB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezl PA;

XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;

XX WPI; 2003-46849/44.

PT Determining the presence or absence of a pathological cell in a patient;

XX useful for diagnosing, prognosing or treating cancer, comprises detecting

XX a nucleic acid in a biological sample.

PS Claim 8; SEQ ID NO 515; 1385bp; English.

XX The invention relates to nucleic acids and proteins (ADN3683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularization syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.

XX SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-135 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271B-14 (1-293) x ADN39197 (1-1260)

QY 1 MetAlaThrAlaArgProProTpmMetTpvValLeuCyAlaLeuIleThrAlaLeu 20
DB 43 ATGGCTACAGACAGACCCCTGATGTGGTCTGTGCTGTGATTCACAGCTTGTCT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAa 40
DB 103 CTGGGGGTACAGAGATGTTCTGCCAATGATGTTTCTGTACCAACCTCTTAAAC 162
QY 41 ThrValProSerGlySerArgHisAspLeuGlyValGlyValAlaGlyValAspAlaArgSer 60
DB 163 ACCGTCCTCTGGGAGCAACAGACCTGGGAGCTGGGGCCGGGAGAAAGCCGCCGTG 222
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProT 80
DB 223 GATGACAGACAGCCGCATCATCATGATCCGACTGGCATATGACACCCGACCTGG 282
QY 81 GlnAlaIleLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValAlaValLeuValHisPro 100
DB 283 CAGGCCGCTGTGTAAAGCCCAACAGCTTACTGGCGGGCGGTGGTGGATCCA 342
QY 101 GlnTPrpLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValArgLeuGlyHis 120
DB 343 CAGTGGCTCTACAGCCGCCCTCAGAGAAAGATTTCAGATCCGTCTCGGCCAC 402
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
DB 403 TACTCCCTGTACCAATTTATGATCTGGGACAGCAATGTTCCAGAGGGGTCAAAATCATC 462
QY 141 ProHisProGlyTyrIserHisProGlyHisIserAsnAspLeuMetLeuIleIysLeuAa 160
DB 463 CCCCACTCTGGTACTCCCACTGGCCACTTAAAGACCTCATGTCTCAAAACGAAAC 522
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 523 AGAAGATTCTGCCCTAAAGATGTCAGACCATCAAGCTCTCTCATGTCTCCCTCT 582
QY 181 AlaGlyThrIleCysLeuValSerGlyTPrpGlyThrThrIysSerProGlnValHisPhe 200
DB 583 GCTGGAGACAAAGTCTGTGTGTGGCTGGGAGCAACAGAGCCCAAGAGCACTTC 642
QY 201 ProIysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlyAspArgCysGluAspAla 220

DB 643 CCTAAGTCTCTCAGAGCTTGAATATCAGCGTGTAAAGTCAAGAAAGGCGAGATGCT 702
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspValArgAspSer 240
DB 703 TACCCGAGACAAATGATATACACATGTTCTGCGCGGTGAACAAAGCAGTGAAGCTTC 762
QY 241 CysGlnGlyAspSerGlyValProValValCysAsnGlySerLeuGlnIlyLeuValSer 260
DB 763 TCCAGAGGATGATTTGGGGGGCTGTGTGTCTCATATGCTCTCGAGGAGACTCGTCTC 822
QY 261 TPrpGlyAspTyrProCysAlaArgProAsnArgProGlyValIlyThrAsnLeuCysIlys 280
DB 823 TGGGAGATTAACCTTGTGCGCGCCCAACAGACCGGGGTGTACACGAACCTTGCAAG 882
QY 281 PheThrIleTyrIleGlnIlyThrIleGlnAlaAsnSer 293
DB 883 TTCACCAAGTGAATCCAGAAACCATCCAGGCCAATCTCC 921

RESULT 6

ID ADR25550
ADR25550 standard; DNA; 1260 BP.

XX AC ADR25550;
XX XX 21-OCT-2004 (first entry)
XX DT 21-OCT-2004 (first entry)
XX XX
XX DE Breast cancer prognosis marker #1411.
XX KW ds; breast cancer; prognosis; gene expression; diagnosis.
XX KW Homo sapiens.
XX OS
XX PN WO2004065545-A2.
XX PD 05-AUG-2004.
XX PF 15-JAN-2004; 2004WO-US001100.
XX PR 15-JAN-2003; 2003US-00342887.
XX PA (ROSE-) ROSETTA INPHARMATICS LLC.
XX PA (NECA-) NETHERLANDS CANCER INST.
XX PI Van't Veer LJ, He Y;
XX DR WPI; 2004-593473/57.
XX PT Classifying a breast cancer patient according to prognosis comprises
XX PT determining the similarity between the level of expression of each of
XX PT five genes in a cell sample taken from patient, to control levels.
XX PS Disclosure; SEQ ID NO 1411; 226bp; English.
XX PS
XX CC The invention relates to a method of classifying a breast cancer patient
XX CC according to prognosis by determining the similarity between the level of
XX CC expression of each of five genes for which markers are listed in the
XX CC specification, in a cell sample taken from the breast cancer patient, to
XX CC control levels of expression for each respective five genes to obtain a
XX CC patient similarity value. The methods are useful for classifying a breast
XX CC cancer patient according to prognosis. Kits and computer program products
XX CC are useful for data analysis using the diagnostic, prognostic and
XX CC statistical methods of the invention. This sequence corresponds to a
XX CC marker used in the method of the invention.

SQ Sequence 1260 BP; 276 A; 395 C; 319 G; 270 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.18e-135 Length: 1260
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0

DB:	13	Gaps:	0
DB:	US-09-936-271B-14 (1-293) x ADR25550 (1-1260)		
OY	1 MetAlaIThrAlaArgProTyrPMeTTrpValLeuCyAlaLeuIleThraIaLeuLeu 20		
Db	43 ATGGCTACAGCAAGACCCCTCGATGTGGGTGTCTGTGCTCGATCAGACCTTCTT 10		
OY	21 LeuGlyValIThrGluHsValLeuAlaAsnAspValSerCyAspAlaProSerAsn 40		
Db	103 CTGGGGGTACAGAGCATGTTCTGGCCAAACAATGATGTTCTGTGACCACTTCTAAC 16		
OY	41 ThrValProSerGlySerAsnGlnAspLeuGlyValAlaGlyValAspAlaArgSer 60		
Db	163 ACCGTGCCCTCTGGAGCAACAGAGACTGGAGCTGGGGCCGGGGAAGACGCCGGTTCG 22		
OY	61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80		
Db	223 GATGACAGACAGACCGGCATCATGATGCCATCTGCATATGACACACCGCCTGG 28		
OY	81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyValAlaLeuValHisPro 100		
Db	283 CAGGCGCGCTGTGCTTAAGGCCCAACACAGCTCTACTGCGGGGGGGTGGTGCATCCA 34		
OY	101 GlnTyrPLeuLeuThrAlaAlaHisCyAlaGlyLeuValaPheArgValaIleGluGlyHis 120		
Db	343 CAGTGGCTGCTCAAGGCGCGCCCTCACTCAGGAMAATAATTTTCAGAGCCGCTTCGGCCAC 40		
OY	121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValaLysSerIle 140		
Db	403 TACTCCCTCTCACCAATTATGATCTGGGACAGATGTTCCAGGGGGGTCAATTCATC 46		
OY	141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160		
Db	463 CCCCAACCTCGGCTACTCCACCTGGCCACTTAAAGACTCATGTGCATCAACTAAC 52		
OY	161 ArgAlaGlyIleArgProThrLysAspValaIleArgProIleAsnValSerSerHisCyProSer 180		
Db	523 AGAAGAAATTCGTCCTAAAGATGTCAGACCCATCAACGCTCTCTCATTTGCCCTCT 58		
OY	181 AlaGlyThrLysCyLeuValSerGlyTyrGlyTyrThrLysSerProGlnValHisPhe 200		
Db	583 GCTGGGACAAAGTCTGTTGGTCTGGCTGGGGACAAACAAAGACCCCAAGTGCACCTTC 64		
OY	201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220		
Db	643 CCTAAGTCTCTCAAGTCTTGAATATCAAGGTGCTAAGTCAAGAAAGTGCAGAGAGCT 70		
OY	221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240		
Db	703 TACCCGAGACAAATGATGATACACCAAGTTCCTGGCCGGGTGACAAAGCAGGTAAAGACTCC 76		
OY	241 CysGlnGlyAspSerGlyGlyProValaCysAsnGlySerLeuGlnGlyLeuValSer 260		
Db	763 TGCACGGGTGATTTCTGGGGGGCTGTGGTCTGCAATGAGCTCCCTGCAGGGAGCTCGTCTC 82		
OY	261 TyrGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCyLys 280		
Db	823 TGGGAGAAATTAACCTTGTGTCCCGGGCCCAAGACCGGGTGTCTACACGAACCTCTGCAG 88		
OY	281 PheThrLysTyrIleGlnGluThrIleGlnAlaAsnSer 293		
Db	883 TTCACCAAGTGGATCCAGGAAACATCCAGGCCAACTCC 921		
RESULT 7			
ID	AAK57990		
XX	AAK57990		
AC	AAK57990;		
XX	AAK57990;		
DT	19-JUL-1999 (first entry)		
DE	Human BS247 specific polynucleotide #8.		

	XX	B5247; detection; diagnosis; breast cancer; atypical hyperplasia;
KM	XX	fibroadenoma; cystic breast disease; gene therapy; ss.
OS	XX	Homo sapiens.
Dn	XX	W09922027-A1.
Pd	XX	06-MAY-1999.
Pf	XX	28-OCT-1998; 98MO-US022906.
Pr	XX	28-OCT-1997; 97US-006343IP.
Xx	XX	28-OCT-1997; 97US-00968838.
PA	(ABBO) ABBOTT LAB.	
PI	Billing-Medel PA, Cohen M, Colpitte TL, Friedman RN, Gordon J;	
Pt	Granados EN, Hodges SC, Klass MR, Kratochvil UD, Russell UC,	
Ps	Strope du, Yu H;	
PP	WI, 1999-312977/26.	
PT	Breast tissue derived cDNA contig and consensus polypeptide sequence.	
PS	Claim 1; Page 106; 112pp; English.	
CC	This sequence is a B5247 specific polynucleotide. The invention relates	
CC	to a method of detecting the presence of a target B5247 polynucleotide,	
CC	especially mRNA, in a test sample. B5247 polynucleotides are derived from	
CC	breast tissue. The polynucleotides, polypeptides or antibodies are useful	
CC	for providing information leading to the detection, diagnosis, staging,	
CC	monitoring, prognosis, in vivo imaging, prevention or treatment,	
CC	determining predisposition to, diseases and conditions of the breast,	
CC	such as breast cancer, atypical hyperplasia, fibroadenoma and cystic	
CC	breast disease. Drug treatment or gene therapy for breast cancer, can be	
CC	based on these identified gene sequences and the efficacy of any	
CC	particular therapy can be monitored. The B5247-derived reagents are	The
CC	advantageous for detection of breast cancer due to their specificity.	
CC	Reagents also provide an alternative, non-surgical diagnostic method	
CC	capable of detecting early stage breast disease, such as cancer	
SQ	Sequence 1302 BP, 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;	
Alignment Scores:		
Pred. No.:	1, 23e-135	Length: 1302
Score:	1608.00	Matches: 293
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	2	Gaps: 0
US-09-936-271B-14 (1-293) x MAX57990 (1-1302)		
Qy	1 MetAlaThrAlaArgProBioTrpMetTrpValIleuCyAlaLeuIlleThrAlaLeu	20
Db	88 ATGGTACAGCAAAACCCTCGATGTGAGTGCTGTGCTGCATCAGACTTCCT	147
Qy	21 LeuGlYValThlGluHisValLeuLAlaAnnsApVAlSerCYaaPAPRProSeAan	40
Db	148 CTGGGGGTACAAGAGATTCTCCCAAATATGTTCCTGTGAACACCCCTCAC	207
Qy	41 ThrValProSerGIYSerAndInAspLeuGluYlaAglyAlaGyluaAPAAlaArgSer	60
Db	208 ACCGTGCCCTCTGGAGNACAACAGAACCTGGAGCTGGGGCCCCGGAGAGGCCGGTGG	267
Qy	61 AspASPserSerserArgllelleAnnlySerAsPCysaPMechIsthnrglnProtpr	80
Db	268 GATGACAGCAGAGCGCATATCATATGATCCAGTCCAGTACGATATGACACACCGCGTGG	327
Qy	81 GluAlAlAlaLeuLeuLuarPProAnglnLeutryCYsgelyalavalLeuvalHIsPro	100
Db	328 CAAGCCGCCTTTGTCTTAAGCCCAACAGCTACTACGCGGGGGGTGTGGTGCAATCCA	387

```
QY 101 GINTPLeuLeuThraAlaAlaHisCyAsrGlyLeuValPheArgValArgLeuGlyHis 120
XX |||
PS 388 CAGTGGCTCTCAGCGCCGCCCTCAGAGAAAGATTTCAGATCCCTCTCGGCCAC 447
DB |||
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
XX |||
PS 448 TACTCCCTGTCAACAGTTTATGAAATCTGGGCGACAGATGTTCCAGAGGGGTCAATCCATC 507
DB |||
QY 141 ProHisProGlyTyrThrHisPProGlyHisSerAspAspLeuMetLeuIleLysLeuAn 160
XX |||
PS 508 CCCACCCCTGGCTACTCCACCCCTGGGCCACTTAACGACCTCAATGCTCAATCAACGTAAC 567
DB |||
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
XX |||
PS 568 AGAAGAAATTCGTCCCTCAAGATGTCAGACCCATCACTCTCTCTCTCTCTCTCTCTCT 627
DB |||
QY 181 AlaGlyThrIlyCyLeuValSerGlyTyrPglYThrThrIlySerProGlnValHisPhe 200
XX |||
PS 628 GCTGGGACAAAGCTTGGTGTCTGGCTGGGAGACCAACAGAGCCCCCAAGTGCCTTC 687
DB |||
QY 201 ProIlyValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlyAsrGlyAspAla 220
XX |||
PS 688 CCTAAGCTCTCCAGTGTGATTGAAATATCAGCGTCTAAGTCAAGAAAGGTGCGAGATGCT 747
DB |||
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIlyAspIlyArgAspSer 240
XX |||
PS 748 TACCCGAGACAGTATGATGACACCAATGTTCTGGCCGGGTGACAAACAGGTAGAGACTCC 807
DB |||
QY 241 CysGlnIlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnIlyLeuValSer 260
XX |||
PS 808 TGCACAGGTGATTCGAGGGGGCCCTGTGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCC 867
DB |||
QY 261 TrpGlyAspThrProCysAlaArgProAsnArgProGlyValIlyThrAsnLeuCysIlyS 280
XX |||
PS 868 TGGGGAGATTAACCTTGTCGCCGCCCAACAGACCGGGTCTCAACGAACTCTCAAG 927
DB |||
QY 281 PheThrIlySerIleGlnIlyThrIleGlnIlyAlaAsnSer 293
XX |||
PS 928 TTCACCAAGTGAATCCAGGAAACATCCAGGCCCACTCC 966
DB |||
RESULT 8
AA57989
ID AA57989 standard; DNA; 1302 BP.
XX
AC AA57989;
XX
DT 19-JUL-1999 (first entry)
XX
DE Human BS247 specific polynucleotide #7.
XX
KM BS247: detection; diagnosis; breast cancer; atypical hyperplasia;
XX fibroadenoma; cystic breast disease; gene therapy; ss.
XX
OS Homo sapiens,
XX
PN WO922027-A1.
XX
PD 06-MAY-1999.
XX
PF 28-OCT-1998; 98WO-US022906.
XX
PR 28-OCT-1997; 97US-0063431P.
XX
PR 28-OCT-1997; 97US-00968836.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
XX Grandos EN, Hodges SC, Klass MR, Kratochvil JD, Russell JC,
XX Stroepe SD, Yu H;
XX
DR WPI, 1999-312977/26.
XX
```

```
PT Breast tissue derived cDNA contig and consensus polypeptide sequence.
XX
XX Claim 1, Page 105; 112pp: English.
XX
CC This sequence is a BS247 specific polynucleotide. The invention relates
CC to a method of detecting the presence of a target BS247 polynucleotide,
CC especially mRNA, in a test sample. BS247 polynucleotides are derived from
CC breast tissue. The polynucleotides, polypeptides or antibodies are useful
CC for providing information leading to the detection, diagnosis, staging,
CC monitoring, prognosis, in vivo imaging, prevention or treatment,
CC determining predisposition to, diseases and conditions of the breast,
CC such as breast cancer, atypical hyperplasia, fibroadenoma and cystic
CC breast disease. Drug treatment or gene therapy for breast cancer, can be
CC based on these identified gene sequences and the efficacy of any
CC particular therapy can be monitored. The BS247-derived reagents are
CC advantageous for detection of breast cancer due to their specificity. The
CC reagents also provide an alternative, non-surgical diagnostic method
CC capable of detecting early stage breast disease, such as cancer
XX
SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,23e-135 Length: 1302
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-936-271B-14 (1-293) x AA57989 (1-1302)
QY 1 MetAlaThrAlaArgProProThrMetTrpValLeuCyAlaLeuIleThrAlaLeuLeu 20
XX |||
PS 88 ATGGCTACAGCAAGAACCCCCCGATGTGGGTGCTGTGTGCTTATACACACCTTGCTT 147
DB |||
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCyAspHisProSerAsn 40
XX |||
PS 148 CTGGGGGTTCACAGAGCATGTTCTCCGCAACATGATGTTTCTGTGACACACCCCTTAC 207
DB |||
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaValAlaValIlyAspSer 60
XX |||
PS 208 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAGAGACCCCGAGTGC 267
DB |||
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
XX |||
PS 268 GATGACAGAGCAGCGCATCATCATGATGATCGATGCAATGCACACCGAGCGTGG 327
DB |||
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
XX |||
PS 328 CAGGCCGCGCTGTGCTTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGGTGCATCCA 387
DB |||
QY 101 GINTPLeuLeuThraAlaAlaHisCyAsrGlyLeuValPheArgValArgLeuGlyHis 120
XX |||
PS 388 CAGTGGCTCTCAGCGCCGCCCTCAGAGAAAGATTTCAGATCCCTCTCGGCCAC 447
DB |||
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIysSerIle 140
XX |||
PS 448 TACTCCCTGTCAACAGTTTATGAAATCTGGGCGACAGATGTTCCAGAGGGGTCAATCCATC 507
DB |||
QY 141 ProHisProGlyTyrThrHisPProGlyHisSerAspAspLeuMetLeuIleLysLeuAn 160
XX |||
PS 508 CCCACCCCTGGCTACTCCACCCCTGGGCCACTTAACGACCTCAATGCTCAATCAACGTAAC 567
DB |||
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
XX |||
PS 568 AGAAGAAATTCGTCCCTCAAGATGTCAGACCCATCACTCTCTCTCTCTCTCTCTCTCT 627
DB |||
QY 181 AlaGlyThrIlyCyLeuValSerGlyTyrPglYThrThrIlySerProGlnValHisPhe 200
XX |||
PS 628 GCTGGGACAAAGTGTGTGTCTGGCTGGGAGACCAACAGAGCCCCCAAGTGCCTTC 687
DB |||
QY 201 ProIlyValLeuGlnCysLeuAsnIleSerValLeuSerGlnIlyAsrGlyAspAla 220
XX |||
```


DB 969 TTCACCAAGTGATCCAGGAAACCATCCAGGCCAAGCTCC 1007
|||||
RESULT 10
ADA40370
ID ADA40370 standard; cDNA, 1381 BP.
XX
XX ADA40370;
XX
XX 20-NOV-2003 (first entry)
XX
DE Human secreted protein encoding cDNA.
XX
XX Human; secreted protein; cancer; hyperproliferative disorder;
XX rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
XX anaemia; allergic reaction; asthma; cardiovascular disorder;
XX wound healing; cytostatic; immunosuppressive; nocotropic; neuroprotective;
XX antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
XX vulnary; cardiac; gene therapy; ss.
XX
XX Homo sapiens.
OS
PN MO2002102993-A2.
PD 27-DEC-2002.
PF 19-MAR-2002; 2002MO-US008123.
PR 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-175238/17.
DR
XX
XX New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.
XX
XX Claim 9; SEQ ID NO 752; 3205bp; English.
XX
XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC treating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,33e-135 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0
US-09-936-271B-14 (1-293) x ADA40370 (1-1381)
QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeuLeu 20
DB 129 ATGGGTACAGCAAGAACCCCGTGGATGTGGGTCTGTGCTCTGATCAACACCTTGGCT 188
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCyAspHisProSerAsn 40
DB 189 CTGGGGGTCAAGACATGTCTCCGCAACATGATGTTTCTGTGACCAACCCCTTAAC 248
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB 249 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAGAGACCCCGGTG 308
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 309 GATGACAGACAGACCGCATCATCATGATCCGATGATATGATCAACCCAGCCGTG 368
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCyGlyAlaValLeuValHisPro 100
DB 369 CAGCCCGCTGTCTTAAGGCCCAACGACTCTGCTGGGGCCGGGTGTGTGATCCA 428
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspArgValAspValPheArgValArgLeuGlyHis 120
DB 429 CAGTGGCTGCTACGCGCCGCACTGCAGAGAAAGATTTCAGATCCGTCCGACC 488
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnIleMetPheGlnGlyValIleSerIle 140
DB 489 TACTCCCTGTACCAAGTTTATATATCTGGGACAGATGTTTCCAGGGGTCAAAATCCATC 548
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuLeuIleValLeuAsn 160
DB 549 CCCCAACCTGGCTACTCCCAACCTGGCCCACTTAACGATCATCTCATCAAACTGAAC 608
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyProSer 180
DB 609 AAGAGATTCGTCCCACTAAGATGTCAGACCCATCAACGTCTCTCATTTGCCCTCT 668
QY 181 AlaGlyThrValCyLeuValSerGlyTyrGlyThrThrValSerProGlnValHisPhe 200
DB 669 GCTGGGACAAATGCTGTGTGTCTGGTGGGAGCAACCAAGCCCAAGTGCATCTC 728
QY 201 ProValValLeuGlnCyLeuAsnIleSerValLeuSerGlnValArgCyGluAspAla 220
DB 729 CCTAAGTCTCTCAATGCTTGAATATCAGCGTCAAGTCAAAAAAGGTCCAGAGATGCT 788
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAlaGlyArgAspSer 240
DB 789 TACCCGAGCAATATGATATCAACATGTTCTGCGCGGAGCAAAAGAGTGAAGCTCC 848
QY 241 CyGlnGlyValAspSerGlyValProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 849 TCCAGAGGTGATCTGGGGGGCTGTGTGTCTGCATATGCTCTCCCTCAGGAGATCTGTCTC 908
QY 261 TrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyIle 280
DB 909 TGGGAGATTAACCTTGTGCGCGGCCCAACAGACGGGGGTCTACAGCAACCTTGCAAG 968
QY 281 PheThrIleTrpIleGlnGluThrIleGlnAlaAsnSer 293
DB 969 TTCACCAAGTGATCCAGGAAACCATCCAGGCCAAGCTCC 1007

RESULT 11
ADA56535
ID ADA56535 standard; DNA; 1381 BP.
XX
XX ADA56535;
XX
DT 20-NOV-2003 (first entry)
XX
DE Gene encoding human secreted protein #257.
XX
XX immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
KM cytosolic; cerebroprotective; neuroprotective; neurotropic;
KM cardiovascular; antiatherosclerotic; gene therapy;
KM human secreted protein; immune disorder; inflammation;
KM respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KM inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KM multiple sclerosis; ischemic brain injury; Parkinson's disease;
KM Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KM triple helix formation; antisense gene therapy; forensic biology; ds;
KM gene.
XX
XX Homo sapiens.
OS
XX
XX WC02002102994-A2.
PN
PD 27-DEC-2002.
XX
XX 19-MAR-2002; 2002WC-US008278.
PF
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI WPI; 2003-167512/16.
DR P-PSDB; ADA57428.
DR
XX
XX New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX
XX
XX Claim 21; SEQ ID NO 724; 1754bp; English.
XX
XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that bind to
CC the polypeptides, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorder, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to a gene encoding one of the polypeptide of the invention. Note: The

CC sequence data for this patent did form part of the printed specification,
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
SQ Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,336-135 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-936-271b-14 (1-293) x ADA56535 (1-1381)
QY 1 MetAlaThrAlaArgProPOTRMEtTTPValLeuCYsAlaLeuIleThraAlaLeu 20
DB 129 ATGGCTACAGCAAGACCCCTGGATGTGGTCTGTCTGTATCACAGCCTTGCTT 188
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 189 CTGGGGGTGCACAGAGCATGTTCTGCGCAACAAATGATGTTCTGTGACCACTCCCTTAC 248
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyValAspAlaArgSer 60
DB 249 ACCGTGCTCTTGGAGCAACAGACCTGGAGCTGGGGCCGGAGAACGCCGGTTCG 308
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 309 GATGACAGCAGCAGCGGCATCATCATGATCCGACGTGGATATGACACCCAGCGGTGG 368
QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB 369 CAGGCCGCGCTGTGTCTTAAGCCCAACCACTCTACTGCGGGGGGTGTGTGCATCCA 428
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyLeuValPheArgValArgLeuGlyHis 120
DB 429 CAGTGGCTGTCAAGCGGCCCACTGCGAGAAAGAACTTTTCAGATGCCGTCTGGGCAC 488
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValYsSerIle 140
DB 489 TACTCCCTGTACACAGTTTATGATCTGGGCACAGATGTTTCCAGGGGGTCAATCATC 548
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuIleValYsLeuAsn 160
DB 549 CCCCACTGGCTACTCCACCTGGGCACCTTAACGACCTCATGCTCATCAAACTGAAC 608
QY 161 ArgArgIleArgProThrTrpAspValArgProIleAsnValSerSerHisCyProSer 180
DB 609 AGAAGATTGCTCCACTAAAGATGTCAGCCCATCAACGTCTCTCTCATATGCTCTCT 668
QY 181 AlaGlyThrYsCyLeuValSerGlyTyrGlyThrThrYsSerProGlnValHisPhe 200
DB 669 GCTGGGACAAAGTCTTGCTGTGTGGTGGGACAAACAAAGCCCCCAAGGCACTTC 728
QY 201 PolyValLeuGlnCYsLeuAsnIleSerValLeuSerGlnYsArgCyGluAspAla 220
DB 729 CTTAAGGCTCTCCAGTGTCTGAATATCAGCTGTAACTGAGAAAGTCCAGGAAGCTT 788
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAspAlaArgAspSer 240
DB 789 TACCCGAGACAGATGATGACACCATGTTCTGGCCGGGTACAAAGAGGTAGAGATCC 848
QY 241 CysGlnIlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 849 TGCAGAGGTATTTGGGGGGCTGTGTGTGCAATGCTCCCTCGACAGGAACTGTGTC 908
QY 261 TrpGlyAspTrpProCyValAspProAsnArgProGlyValYsThrAsnLeuCyGly 280
DB 909 TGGGGAATTAACCTTGTGCGGCCCAACAGACCGGGTGTCTTAACGAACCTTGTGAG 968
QY 281 PheThrYsTrpIleGlnGlnIleThrIleGlnAlaAsnSer 293

XX Human ovarian cancer-related tumour marker kallikrein 5 (KLK5) DNA 1.
DE kallikrein 5, tumour marker; ovarian cancer;
KM epithelial ovarian carcinoma; human; serine protease; chromosome 19q13.4;
KW KLK5; ds; gene.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 170..1051
FT /*csg= a
FT /product= "Human ovarian cancer-related tumour marker
FT kallikrein 5 (hks) protein"
XX
XX MO2004075713.A2.
XX
XX 10-SBP-2004;
XX
XX 26-FEB-2004; 2004MO-CA000281.
XX
XX 26-FEB-2003; 2003US-0450406P.
XX
XX (MOUN) MOUNT SINAI HOSPITAL.
XX
XX Diamandis EP.
XX
XX WPI; 2004-6d1815/64.
XX P-PSDB; ADR72873.
XX
XX PT Kallikrein markers detection method for detecting ovarian cancer in
PT patient, involves detecting kallikrein markers and CA125 in sample
PT obtained from patient, and comparing detected amounts with standard
PT amounts.
XX
XX
XX Example 2; SEQ ID NO 4; 102bp; English.
XX
XX The invention relates to a novel method for detecting a plurality of
CC kallikrein markers associated with ovarian cancer. The method comprises
CC obtaining a sample from a patient and detecting in the sample a plurality
CC of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
CC the kallikrein markers are selected from the group consisting of
CC kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
CC kallikrein 11. The detected amounts of the kallikrein markers are
CC compared with standard amounts. The method of the invention may be useful
CC for detecting kallikrein markers associated with ovarian cancer in a
CC patient and thus for detecting ovarian cancer, particularly epithelial
CC ovarian carcinoma. The current sequence is that of the human ovarian
CC cancer-related tumour marker kallikrein 5 (KLK5) DNA 1 of the invention
CC which encodes a secreted serine protease and is located at chromosome
CC 19q13.4.
XX
XX
SQ Sequence 1387 BP; 290 A; 437 C; 351 G; 309 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
Pred. No.: 1,33e-135 Length: 1387
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-936-271B-14 (1-293) x ADR72874 (1-1387)
QY 1 MetAlaThrAlaArgProCotTrpMetTrpValLeuCYsAlaLeuIleThrAlaLeu 20
170 ATGGCTACACCAAGACCCCTGGATGGGTGCTCTGTGCTGATCACAGCCTTCTT 229
Db 170 ATGGCTACACCAAGACCCCTGGATGGGTGCTCTGTGCTGATCACAGCCTTCTT 229
QY 21 LeuGlyValThrGluHleValLeuAlaAsnAspValSerCYsAspHisProSerAsn 40
230 CTGGGGGTCAACAGAGATGTTCTGCCAACATATATATTTCTGTGACCAACCCCTTAC 289
Db 230 CTGGGGGTCAACAGAGATGTTCTGCCAACATATATATTTCTGTGACCAACCCCTTAC 289
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyAlaGlyIuAspAlaArgSer 60

Db 290 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAAGACCCGGCTCG 349
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCYsAspMetHisThrGlnProCTP 80
Db 350 GATGACAGACAGACCGCATCATCAATGATCGAGCTGCGATGTGACACCCAGCCGTGG 409
QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCYsGlyValAlaValLeuValHisPro 100
Db 410 CAGGCGCGCTGTGCTTAAGCCACACAGCTTACTGCGGGCGGTGGTCATCA 469
QY 101 GlnTrpLeuLeuThrAlaAlaHisCYsArgLYLeValPheArgValArgLeuGlyHis 120
Db 470 CAGTGGCTGCTCAACGCGCCGCTGACAGAAAGAACTTTCAGAGTCCGCTCGGAC 529
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 530 TACTCCCTGTCAACAGATTATGATCTGGGACACAGATGTTCCAGGGGGTCAAAATCATC 589
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspMetLeuIleValLeuAsn 160
Db 590 CCCACCCCTGCTACTCCACCCCTGGCCTTACGACCTCAATGCTCATCAACCTAAC 649
QY 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCYsProSer 180
Db 650 AGAAGATTGCTGCTCAATAGATGTCAGACCCATCAAGCTCTCTCTCATTTGCTCT 709
QY 181 AlaGlyThrIlyCYsLeuValSerGlyTyrGlyThrThrIlySerProGlnValHisPhe 200
Db 710 GCTGGGCAAAAGTCTGTGCTGGCTGGGGGACAAACCAAGAGCCCAAGTGCATCTT 769
QY 201 ProLYsValLeuGlnCYsLeuAsnIleSerValLeuSerGlnIlyAspArgCYsGlyAspAla 220
Db 770 CTTAAGGCTCTCAAGCTTGAATATCAAGGCTCTAAGTCAAAAGGTGCGAGATGCT 829
QY 221 TyrProArgGlnIleAspAspThrMetPheCYsAlaGlyAspIlyValArgAspSer 240
Db 830 TACCCAGACAGATATATGACATCATTTCTGGCCGGTGCACAAAGAGGTATAGATCC 889
QY 241 CysGlnGlyAspSerGlyGlyProValValCYsAsnGlySerLeuGlnGlyLeuValSer 260
Db 890 TGCAGAGGTATCTGGGGGGGCTGTGTGCTGCAATGGCTCCCTGCAGGGAGCTGTGTC 949
QY 261 TrpGlyAspIlyProCYsAlaArgProAsnArgProGlyValIlyThrAsnLeuCYsLys 280
Db 950 TGGGAGATTACCTCTGTGCTGGCCCAACAGACCGGGTGTCAACGAACCTCTGACAG 1009
QY 281 PheThrIlySerTrpIleGlnIuThrIleGlnAlaAsnSer 293
Db 1010 TTACCAAGTGAATCCAGGAACCATTCAGGCCAATCC 1048
RESULT 14
ADA40367
ID ADA40367 standard; cDNA; 1490 BP.
XX
XX ADA40367;
XX
XX 20-NOV-2003 (first entry)
XX
XX Human secreted protein encoding cDNA.
XX
XX Human; secreted protein; cancer; hyperproliferative disorder;
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
KW anaemia; allergic reaction; asthma; cardiovascular disorder;
KW wound healing; cytoskeletal; immunosuppressive; nootropic; neuroprotective;
KW antiviral; anti-allergic; hepatotropic; antidiabetic; anti-inflammatory;
KW vulnery; cardiac; gene therapy; ss.
XX
XX Homo sapiens.
XX
XX WO2002102993-A2.
XX
XX 27-DEC-2002.

XX 19-MAR-2002; 2002WC-US008123.
PF
XX 21-MAR-2001; 2001US-0277340P.
PR 19-UTL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0331287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
XX WPI; 2003-175238/17.
DR
XX
PT New human secreted proteins and nucleic acid molecules, useful for
PT preparing a diagnostic or pharmaceutical composition for diagnosing,
PT preventing or treating cancer or other hyperproliferative disorder,
PT asthma, allergies or AIDS.
XX
XX Claim 9; SEQ ID NO 749; 3205bp; English.
XX
XX The invention relates to novel genes ADA39629-ADA40565 and proteins
CC ADA40566-ADA41501 for human secreted proteins, useful for preventing,
CC creating or ameliorating medical conditions e.g. by protein or gene
CC therapy. The polypeptides, nucleic acid molecules, antibodies or their
CC fragments, and agonists or antagonists that bind to the polypeptide are
CC useful for preparing a diagnostic or pharmaceutical composition for
CC diagnosing or treating cancer or other hyperproliferative disorder. The
CC polypeptides and nucleic acid molecules are also useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating cancer
CC or other hyperproliferative disorders including neoplasms, autoimmune
CC disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematological disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for chromosome identification, radiation hybrid mapping or long-
CC range restriction mapping, as molecular weight markers, or as
CC hybridization or diagnostic probes. The polypeptides and antibodies are
CC useful for providing immunological probes for differential identification
CC of the tissues immunohistochemistry assays. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1490 BP; 323 A; 486 C; 373 G; 307 T; 0 U; 1 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,46e-135 Length: 1490
XX Score: 1608.00 Matches: 293
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 8 Gaps: 0
XX
XX US-09-936-271B-14 (1-293) x ADA40367 (1-1490)
XX
XX 1 MetAlaThrAlaArgProPheTyrValLeuCyAlaLeuIleThrAlaLeuLeu 20
DB 182 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTCTGCTCTGATCACAGCTTGCTT 241
XX
XX 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 242 CTGGGGGTACAGAGCATGTTCTCGCAACATGATGTTCTGTGTGCCACCCCTTAAAC 301
XX
XX 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSer 60
DB 302 ACCGTGCCCTCTGGAGCAACAGACCTGGAGAGCTGGGGCCGGGGAAGACGCCCGGTGCG 361

QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80
DB 362 GATTGACAGCAGCAGCGCATCATCATGATCGACTGGATATGACACCCAGCCGTGG 421
XX
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyAsGlyAlaValLeuValHisPro 100
DB 422 CAGGCCGCGCTGTGTCTAAGGCCCAACGAGCTCTACTGGGGCGGTGTGTGTGATCA 481
XX
QY 101 GlnTyrPheLeuThrAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
DB 482 CAGTGGCTCTACCGCGCCGACCTGACAGAGAAAGATTTCAGATCCGCTCGGCCAC 541
XX
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 542 TACTCCCTGTACCAAGTTTATGATTCGGGACGAGATGTTCCAGGGGTTCAATTCATTC 601
XX
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValSerLeu 160
DB 602 CCCACCCCTGGCTACTCCACCCCTGGCCACTTAAAGACCTCATCTCAATCAACTGAA 661
XX
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerHisCyProSer 180
DB 662 AAGAGATTCGTCCTCACTAAGATGCAACCATCAACGCTCTCTCATTTGCTCTT 721
XX
QY 181 AlaGlyThrIleCyAlaValSerGlyTyrGlyThrThrIleYsserProGlnValHisPhe 200
DB 722 GCTGGGACAAATGCTGTGGTGTCTGGCTGGGGGACACCAAGAGCCCAAGTGCATCT 781
XX
QY 201 ProIleValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnIleValArgCyGluValPhe 220
DB 782 CCTAAGTCTCTCAGTGTGATTCAGCGCTCAAGTCAAGTCAAGTCAAGTCAAGTCT 841
XX
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspIleValArgAspSer 240
DB 842 TACCGAGACATATGATATCACCATTTCTCCCGCGGTGACAAAGCAGTGAACATCTC 901
XX
QY 241 CysGlnGlyAspSerGlyValProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 902 TCCAGGGGTGATTCCTGGGGGCTGTGGTGTCTCAATGCTCCCTCAGGAGCATCTGTTC 961
XX
QY 261 TrpGlyAspTyrProCyAlaArgProAsnAspProGlyValTyrThrAsnLeuCyValys 280
DB 962 TGGGAGATTTACCTTGTGCTCCGCGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAG 1021
XX
QY 281 PheThrIleTyrIleGlnGlnIleThrIleGlnAlaAsnSer 293
DB 1022 TTCACCAAGTGGATCCAGGAACCATTCAGGCCCACTCC 1060
XX
XX RESULT 15
XX ADA56532
XX ID ADA56532 standard; DNA; 1490 BP.
XX
XX AC ADA56532;
XX
XX DT 20-NOV-2003 (first entry)
XX
XX DE Gene encoding human secreted protein #257.
XX
XX KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
KW cytostatic; cerebroprotective; neuroprotective; nocitropic;
KW cardiovascular; antiarteriosclerotic; gene therapy;
KW human secreted protein; immune disorder; inflammation;
KW respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
KW inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
KW triple helix formation; antisense gene therapy; forensic biology; ds;
XX gene.
XX
XX OS Homo sapiens.
XX
XX XX
XX PN W02002102994-A2.
XX

PD 27-DEC-2002.
XX
XX 19-MAR-2002; 2002W0-US008278.
XX
XX 21-MAR-2001; 2001US-0277340P.
PR 19-JUL-2001; 2001US-0306171P.
PR 13-NOV-2001; 2001US-0311287P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
PI
PI; 2003-167512/16.
DR P-PSDB; ADA57425.
XX
XX
PT New human secreted polypeptides and polynucleotides, useful for
PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
PT conditions, respiratory disorders, cancers, CNS disorders, or
PT neurodegenerative disorders.
XX
XX
PS Claim 21; SEQ ID NO 721; 1754bp; English.
XX
XX The invention relates to 592 new human secreted polypeptides useful for
CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
CC conditions, respiratory disorders, cancers, CNS disorders, or
CC neurodegenerative disorders, or polypeptides comprising an amino acid
CC sequence at least 95% identical to the new sequences. The polypeptides,
CC antibodies or antibody fragments that bind to the polypeptides, nucleic
CC acids encoding the polypeptides, agonists or antagonists that binds to
CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
CC compositions for diagnosing, treating or preventing an e.g. immune
CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
CC (e.g. multiple sclerosis or ischemic brain injury), neurodegenerative
CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
CC polynucleotides are useful for chromosome identification, chromosome
CC mapping, for controlling gene expression through triple helix formation
CC or antisense DNA or RNA, in gene therapy, for identifying individuals
CC from minute biological samples, in forensic biology, and as hybridization
CC probes. The polypeptides are useful for as molecular weight markers on
CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
CC gels, to raise antibodies, for testing biological activities, and for
CC treating or preventing neural disorders, immune system disorders,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative and/or cancerous diseases. This sequence corresponds
CC to a gene encoding one of the polypeptide of the invention. Note: The
CC sequence data for this patent did form part of the printed specification,
CC but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 1490 BP; 323 A; 486 C; 373 G; 307 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 1 46e-135 Length: 1490
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-936-271b-14 (1-293) x ADA56532 (1-1490)
QY 1 MetAlaThrAlaArgProTyrMetTyrValLeuCysAlaLeuIleThrAlaLeu 20
Db 182 ATGGCTACAGAAACACCCCTGATGTGGTGTCTGTGTGATCAACAGCTTGTCTT 241
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 242 CTGGGGGTCAAGAGCATGTTCGCCAACATGATGTTCCTGTGACCAACCCCTTAC 301
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyIleuAspAlaArgSer 60

Db 302 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAAGACCCCGCTCG 361
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80
Db 362 GATGACAGACAGACCGGCATCATCAATGATTCGAGCTGAGCATATGACACCCGCGTGG 421
QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCysGlyValValLeuValHisPro 100
Db 422 CAGCGCGCTGTGTGCTAAGCCCAACAGCTTACTGCGGGCGGTGTGCTCA 481
QY 101 GlnTyrLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 482 CAGTGGCTGTCAAGCGCGCCCACTCGAGAAABAAGTTTTCAGAGTCGCTCGGCGAC 541
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerTie 140
Db 542 TACTCCCTGCACAGATTATGATTCGGGACAGATGTTCCAGGGGGTCAATCCATC 601
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspLeuMetLeuIleLysLeuAsn 160
Db 602 CCCACCCCTGGCTACTCCACCTGGCCACTTACGACCTCATGCTCATCAACATGAAC 661
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCysProSer 180
Db 662 AGAAGAAATGTCTCCACTAAGATGTCAGACCATCAAGCTCTCTCATTTGCTCT 721
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db 722 GTGGGACAAAGTCTGGGTGTGGCTGGGGGACCAACCAAGAGCCCAAGTGCATTC 781
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 782 CTTAAGTCTCTCAAGCTTGAATATCAAGCTTAAGTCAAGAAAGTCCGAGATGCT 841
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 842 TACCCAGACAGATATGATGACCATGTTCTGCGCGCGTACAAAGAGATAGATCC 901
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 902 TGCACAGGATATCTGGGGGGCTGTGGTGTGCATATGCTCCCTGCAGGAGATCGTGTCC 961
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 962 TGGAGAAATTAACCTTGTGCTCCGACCAAGACCGGGTGTTCACAGAACTTGTGACG 1021
QY 281 PheThrLysTrpIleGlnGlnIleThrIleGlnAlaAsnSer 293
Db 1022 TTCACCAAGTGAATCCAGGAACCATCCAGGCCACTCC 1060

Search completed: February 26, 2005, 12:58:47
Job time : 617 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:32:30 ; Search time 204 Seconds

(without alignments)
2350.143 Million cell updates/sec

Title: US-09-936-271B-14
Perfect score: 1608
Sequence: 1 MATARPPMVMVLCALITALL.....VYTNLCKFTKMIQTIQANS 293

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0936271/runat_23022005_130127_27287/app_query.fasta_1.455
-DB=Issued_Patents_NA -OPMT=faeap -SUFFIX=rnl -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonsum62 -TRANS=human40.ccd
-LIST=45 -DOCLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=p2n -NOBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USPR=US0936271.QCGN 1.1.66@runat_23022005_130127_27287 -NCPV=6 -ICPV=3
-NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1608	100.0	1499	4 US-09-509-908-1	Sequence 1, Appl1
2	1595	99.2	1476	2 US-08-824-874-2	Sequence 2, Appl1
3	1595	99.2	1476	3 US-09-210-084-2	Sequence 2, Appl1
4	1595	99.2	1476	4 US-09-764-762-2	Sequence 2, Appl1
5	1479	92.0	1143	4 US-09-949-016-1841	Sequence 1841, Ap
6	1403	87.3	1504	3 US-09-280-116-1	Sequence 1, Appl1
7	1002	62.3	10818	4 US-09-949-016-13583	Sequence 13583, A
8	725	45.1	735	3 US-09-602-877A-94	Sequence 94, Appl1
9	671	41.7	765	3 US-09-439-313-524	Sequence 524, App
10	671	41.7	765	4 US-09-636-215-524	Sequence 524, App
11	671	41.7	765	4 US-09-685-166A-524	Sequence 524, App
12	671	41.7	765	4 US-09-679-426-524	Sequence 524, App

13	671	41.7	765	4 US-09-759-143-524	Sequence 524, App
14	671	41.7	765	4 US-09-651-236-524	Sequence 524, App
15	654.5	40.7	1052	3 US-09-386-642-10	Sequence 10, Appl
16	652.5	40.6	1146	4 US-09-205-258-247	Sequence 247, App
17	652.5	40.6	1166	3 US-08-944-483-7	Sequence 7, Appl1
18	650.5	40.5	833	2 US-08-790-137-2	Sequence 2, Appl1
19	650.5	40.5	1192	3 US-08-944-483-8	Sequence 8, Appl1
20	650.5	40.5	1292	4 US-09-205-258-189	Sequence 189, App
21	650.5	40.5	1314	3 US-09-025-059-2	Sequence 2, Appl1
22	641.5	39.9	1343	4 US-09-618-259-72	Sequence 72, Appl1
23	641.5	39.9	1360	4 US-09-618-259-72	Sequence 6, Appl1
24	640	39.8	1350	4 US-09-636-215-616	Sequence 616, App
25	640	39.8	1350	4 US-09-685-166A-616	Sequence 616, App
26	640	39.8	1350	4 US-09-679-426-616	Sequence 616, App
27	640	39.8	1350	4 US-09-759-143-616	Sequence 616, App
28	639.5	39.8	1350	4 US-09-651-236-616	Sequence 616, App
29	639.5	39.8	944	3 US-09-008-271A-19	Sequence 19, Appl1
30	638.5	39.7	944	3 US-09-070-526-1	Sequence 1, Appl1
31	638.5	39.7	1253	4 US-09-949-016-2295	Sequence 2295, Ap
32	638	39.7	871	1 US-08-744-026-2	Sequence 2, Appl1
33	638	39.7	871	2 US-09-102-732-2	Sequence 2, Appl1
34	638	39.7	871	3 US-09-261-767-2	Sequence 2, Appl1
35	638	39.7	871	3 US-08-969-987-7	Sequence 7, Appl1
36	638	39.7	1248	3 US-09-020-956-171	Sequence 171, App
37	638	39.7	1248	3 US-09-030-607-171	Sequence 171, App
38	638	39.7	1248	3 US-09-439-313-171	Sequence 171, App
39	638	39.7	1248	3 US-09-352-616A-171	Sequence 171, App
40	638	39.7	1248	3 US-09-232-145A-171	Sequence 171, App
41	638	39.7	1248	4 US-09-159-812-171	Sequence 171, App
42	638	39.7	1248	4 US-09-636-215-171	Sequence 171, App
43	638	39.7	1248	4 US-09-685-166A-171	Sequence 171, App
44	638	39.7	1248	4 US-09-115-453-171	Sequence 171, App
45	638	39.7	1248	4 US-09-688-489-171	Sequence 171, App

ALIGNMENTS

RESULT 1

US-09-509-908-1
Sequence 1, Application US/09509908

Patent No. 6589770
GENERAL INFORMATION:

APPLICANT: The Procter & Gamble Company, N/A N/A
TITLE OF INVENTION: A Protease

NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: T. David Reed
STREET: 5299 Spring Grove Avenue

CITY: Cincinnati
STATE: Ohio

COUNTRY: USA
ZIP: 45217-1087

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/509,908
FILING DATE: 28-Feb-2002

CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:

NAME: Reed, T David
REGISTRATION NUMBER: 32,931

REFERENCE/DOCKET NUMBER: AA-264P
TELECOMMUNICATION INFORMATION:

TELEPHONE: 513-627-7025
TELEFAX: 513-627-6333

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1499 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

/	TOPOLOGY: linear	
/	MOLECULE TYPE: DNA (genomic)	
/	FEATURE:	
/	NAME/KEY: CDS	
/	LOCATION: 291..1172	
/	FEATURE:	
/	NAME/KEY: mat_peptide	
/	LOCATION: 489..1172	
/	FEATURE:	
/	NAME/KEY: 5'UTR	
/	LOCATION: 1..220	
/	FEATURE:	
/	NAME/KEY: 3'UTR	
/	LOCATION: 1173..1499	
/	SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
/	US-09-509-908-1	
Alignment Scores:		
Pred. No.:	6.98e-164	Length: 1499
Score:	1608.00	Matches: 293
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	4	Gaps: 0
US-09-936-271B-14, (1-293) x US-09-509-908-1 (1-1499)		
QY	1 MetValThrAlaArgProPOTpmETTpValLeuCySaLaLeuLeIeThrAlaLeuLeu	20
Db	291 ATGGCTACACAGAACCCCTCGATGTGGTGCTCTGTGCTGTATACAGCCTTGCTT	350
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCySaSPHisProSerAsn	40
Db	351 CTGGGGGTCAACAGCATGTTCTCGCAACATGATGTTTCTGTGACCACTCTTAC	410
QY	41 ThrValProSerGlySerAnginaSPleuGlyValAGlyValAGlyValAspAlaArgSer	60
Db	411 ACCGTGCTCTTGGAGAACCAAGACCTGGAGCTGGGCGGGAGACGCCGGGTGG	470
QY	61 AspAspSerSerSerArgIleIleAnglySerAspCySaSPMetHisThrGlnProTrp	80
Db	471 GATGACAGCAGACGCCGATCATCATGATGATCCGATCGATATGACACCCACCGTGG	530
QY	81 GluAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCyGlyValAlaValLeuValHisPro	100
Db	531 CAGGCGGCGGTGTGCTTAAGCCCAACGCTTACTGGGGGCGGTGTGTGATCCA	590
QY	101 GlnTrpLeuLeuThrAlaAlaHisCySaArglySLysValPheArgValArgLeuGlyHis	120
Db	591 CAGTGGGTGCTCAAGGCGCGCCACTGCTCAAGAAAGAAATTTTCAAGTCCGTGCGCAC	650
QY	121 TyrSerLeuSerProValTyrGluSerGlyGlnImetPheGlnGlyValLysSerIle	140
Db	651 TACTCCCTGTCAACAGATTATGATATGTGGCGAGCATGTTCCAGGGGGTCAAAATTCATC	710
QY	141 ProHisProGlyTyrIleSerHisASPProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
Db	711 CCCCACTCTGGTACTCTCCACCTGTGCCACTTAAAGAACCTCAGTCTATATAACTAAC	770
QY	161 ArgArgIleArgProThrThrLysAspValArgProIleAsnValSerSerHisCySaProSer	180
Db	771 AGAAGATTGTCCCACTAAAGATGCAGACCCATCAAGTCTCTCTCATTTGCTCTCT	830
QY	181 AlaGlyThrIleArgLysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe	200
Db	831 GCTGGGCAAAAGGCTTTGGTGTGTGGCTGGGGGAAACCAAGAGGCCCAAGTCACTTC	890
QY	201 ProLysValLeuGlnCySaLeuAsnIleSerValLleuSerGlnLysArgCyGlyAspAla	220
Db	891 CTTAAGTCTCTCAAGTCTTGAATATACACGTGCTTAAGTCAGAAAAGTGGAGAGTCT	950
QY	221 TyrProArgGlnIleAspAspThrMetPheCySaIaGlyAspLysValAGlyArgAspSer	240

```

Db      951  TACCCGAGACAGATAGATGACACCATGTTCTGCGCCGCTGACAAAGCAGGTAGAGACTCC 1010
Oy      241  CyGgInglYasBsrGlyglYProValYalCYaBsnGlySerLeuGInglYleuValSer 260
Db      1011  TGCACAGGTATATTCCTGGGGGGCCCTGTGCTGTGCATATGCTCTCCCTGCAGGAGACTGTGTCC 1070
Oy      261  TrpGlyAspTrpProCYaAlaArgProAsnArgProGlyValYrThrAsnLeuCYlys 280
Db      1071  TGGGGAGATTACCTTTGTGCCCGGCCCAACAGACGGGTGTCTACAGAACTCTGCAG 1130
Oy      281  PheThrIysTrpIleGIngluThrIleGlnAlaAsnSer 293
Db      1131  TTCACCAAGTGATGCAGAAACCATCCAGGCCCACTCC 1169

RESULT 2
US-08-824-874--2
/ Sequence 2, Application US/08824874
/ Patent No. 5962300
/ GENERAL INFORMATION:
/ APPLICANT: Hillman, Jennifer L.
/ APPLICANT: Lal, Preeti
/ TITLE OF INVENTION: NOVEL KALLIKREIN
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/824,874
/ FILING DATE: Filed Herewith
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1476 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KERNOT02
/ CLONE: 820694
/ US-08-824-874-2

Alignment Scores:
Pred. No.:      1,73e-162      Length:      1476
Score:          1595.00      Matches:      293
Percent Similarity: 99.66%      Conservative: 0
Best Local Similarity: 99.66%      Mismatches:  0
Query Match:     99.19%      Indels:      1
DB:              2          Gaps:              0

US-09-936-271B-14 (1-293) x US-08-824-874-2 (1-1476)
Oy      1  MetAlaThrAlaArgProPOTrMetTrpValLeuCYaAlaLeuIleThrAlaLeu 20
Db      277  ATGGCTACACACAGACCCCTCGATGTGGGTCTCTGTGCTGTGATCAACAGCCTTGCTT 336

```


QY 21 LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerAsn 40
DB 337 CTGGGGGGGACAGAGCAATGTTCTGCCAACAATATGTTCTGTGACACCCCTCTAAC 396
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyValAspAlaArgSer 60
DB 397 ACCGTGCTCTGTGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAGAGCGCCGGTGG 456
QY 61 AspAspSerSerSerArgGlyLeuAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 457 GATGACAGCAGCAGCCCATCATCAATGATCCAGCTGCGGATGATGACACCCACCCCTGG 516
QY 81 -GlnAlaAlaLeuLeuLeuLeuArgProAsnGlnLeuTrpCysGlyValAlaValLeuValHisPr 100
DB 517 GCAGGCGCGCTGTGTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGTGATCC 576
QY 100 OGIntPLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
DB 577 ACAGTGGCTGTCTACCGGCCCGCCCTGACAGAGAAAGTTTTCAGAGTCCGCTCGGCCA 636
QY 120 sTySerLeuSerProValTyrgLysGlyGlnGlnMetPheGlnGlyValLysSer11 140
DB 637 CTACTCCCTGTACACAGATTATGATCTGGGACAGATGTTCCAGGGGGTCAAAATCAT 696
QY 140 eProHisProGlyTyrsSerHisProGlyHisSerAspLeuMetLeuLysValLeuAs 160
DB 697 CCCCACCTGGCTACTCCACCTGGCCACTTCAAGCACTCATGTCTCATCAAACTGAA 756
QY 160 nArgArgLysArgProThrLysAspValArgProLysAsnValSerSerHisCysProse 180
DB 757 CAGAGAAATTCCTCCCACTAAAGATGTCAAGCCATCACTCTCTCTACTTCTCTCCTC 816
QY 180 rAlaGlyThrLysCysLeuValSerGlyTyrgLysThrLysSerProGlnValHisPh 200
DB 817 TGCTGGGACAAAGCTGTGTGTCTGGGCTGGGACCAACCAAGCCGCCCAAGTGCACTT 876
QY 200 eProLysValLeuGlnCysLeuAsnLysSerValLeuSerGlnLysArgCysGlyAspAl 220
DB 877 CCCCAGGCTCTCTCAGTGTGTAATATCAGGCTCTAAGTCAAGAAAGTCCGAGATGC 936
QY 220 sTyProArgGlnLeuAspAspThrMetPheCysAlaGlyAspLysValGlyArgAspSe 240
DB 937 TTACCCAGACAGATAGTATGACACATCTTCTGGCGCGGTACAAAGCAGGTAGAGCTC 996
QY 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
DB 997 CTGCCAGGGTGATTTCTGGGGGGCTGTGTCTGCAATGGCTCCTGCAAGGAGCTCGTGC 1056
QY 260 rTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrgLysAsnLeuCysLy 280
DB 1057 CTGGGGAGATTATCTGTGTGTGCTGGCCCAACAGACCGGGGTCTTCAACGAACCTTGTCAA 1116
QY 280 sPheThrLysTrpLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 293
DB 1117 GTTACCAAGTGGATCCAGGAACCATTCAGGCCCACTCC 1156

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/210,084
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/824,874
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0252 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-845-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1476 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERNOT02
CLONE: 820694
US-09-210-084-2
Alignment Scores:
Pred. No.: 1,73e-162 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 3 Gaps: 0
US-09-936-271b-14 (1-293) x US-09-210-084-2 (1-1476)
QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuLysThrAlaLeuLeu 20
DB 277 ATGGCTAAGCAGCAAGACCCCTGGATGTGGGTCTGTGCTCTGATCAACAGCTTGCTT 336
QY 21 LeuGlyValThrGluHisValLeuAlaAspAspValSerCysAspHisProSerAsn 40
DB 337 CTGGGGGTCAACAGCATGTTCTGCCCAACATGATGTTCTGTGACCAACCCCTCTAAC 396
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyValAspAlaArgSer 60
DB 397 ACCGTGCTCTGTGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAGAGCGCCGGTGG 456
QY 61 AspAspSerSerSerArgGlyLeuAsnGlySerAspCysAspMetHisThrGlnProTrp 80
DB 457 GATGACAGCAGCAGCCCATCATCAATGATCCAGCTGCGGATGATGACACCCACCCCTGG 516
QY 81 -GlnAlaAlaLeuLeuLeuLeuArgProAsnGlnLeuTrpCysGlyValAlaValLeuValHisPr 100
DB 517 GCAGGCGCGCTGTGTGTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGTGATCC 576
QY 100 OGIntPLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
DB 577 ACAGTGGCTGTCTACCGGCCCGCCCTGACAGAGAAAGTTTTCAGAGTCCGCTCGGCCA 636
QY 120 sTySerLeuSerProValTyrgLysGlyGlnGlnMetPheGlnGlyValLysSer11 140
DB 637 CTACTCCCTGTACACAGATTATGATCTGGGACAGATGTTCCAGGGGGTCAAAATCAT 696
QY 140 eProHisProGlyTyrsSerHisProGlyHisSerAspLeuMetLeuLysValLeuAs 160
DB 697 CCCCACCTGGCTACTCCACCTGGCCACTTCAAGCACTCATGTCTCATCAAACTGAA 756
QY 160 nArgArgLysArgProThrLysAspValArgProLysAsnValSerSerHisCysProse 180

Db	757	CAGAAAGATTGCTCCCACTAAAGATGTAGACCCATCAAGCTCTCTCATTTGCTCTC	816
Qy	180	ralaGlyThrIysCysLeuValSerGlyTyrGlyThrThrIysSerProGlnValHisph	200
Db	817	TGCTGGGCAAAAGTGCTTGCTGCTCCGCTGGGGGGCAACCAAGAGCCCCCAAGTGCACTT	876
Qy	200	eProIysValIeugInCysLeuAsnHisSerValIeusSerGlnIysArgCysGlyaaPAl	220
Db	877	CCCTTAGTGCTCCCAAGTGCTTGATATATACGGCTGTCTAGTACGAAAGAGTCCGAGAGATCC	936
Qy	220	aTyreProArgGlnIleaaPAspThrMetPheCysAlaGlyAspIysAlaGlyArgaspSe	240
Db	937	TTACCCGAGACAGATAGATGACACCATGTTCTGGCCCGGTACCAAGCAGGTAGAGACTC	996
Qy	240	rCyGlnIlyAspSerGlyIyGlyProValValCysAsnGlySerIeugInGlyIleuValSe	260
Db	997	CTGCCAGGATATTTGGGGGGCCCTGTGTCTGCAATGAGCTCCCTGCAGGGAGACTCGGTCTC	1056
Qy	260	rTyrGlyAspTyrProCysAlaArgProAsnArgProGlyIyValTyrThrAsnLeuCysIy	280
Db	1057	CTGGGGAGATTACCTCTTGCTGGCCGGCCCAACGAGACGGGTGTCTTACAGAACCTCTGTGAA	1116
Qy	280	sPheThrIyStrIleGlnInGlyThrIleGlnAlaAsnSer	293
Db	1117	GTTCAACCAAGTGATCCAGAAACCATCTCAGGCCAACTCC	1156

RESULT 4
 US-09-764-762-2, Application US/09764762
 Patent No. 6472195
 GENERAL INFORMATION:
 APPLICANT: Hillman, Jennifer L.
 Inventor:
 TITLE OF INVENTION: NOVEL KALLIKREIN
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STEEER: 3174 Porter Drive
 City: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/764,762
 FILING DATE: 16-Jan-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/210,084
 FILING DATE: <Unknown>
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0252 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1476 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: KERANOT02
 CLONE: 820594
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-764-762-2

Alignment Scores:	
Pred. No.:	1,736-162
Score:	1595.00
Percent Similarity:	99.68%
Best Local Similarity:	99.68%
Query Match:	99.19%
DB:	4
	0
Length:	1476
Matches:	293
Conservative:	0
Mismatch:	0
Indels:	1
Gaps:	0

US-09-936-271B-14 (1-293) X US-09-764-762-2 (1-1476)

[illegible]

QY	81	GlnAlaAlaIleuLeuLeuAaGrProArgGlnIleuTyrCYsGlyAlaValIleuValHisPrr	100
Db	517	GCAGGCCCGCGTGTTCAGGCGCCCAACACACTCTACTCGCGGGCGCGTGTGGTCATCC	576
QY	100	oGlnTrpLeuLeuThrAlaAlaHisCYsArgLYsIleValPheArgValAAGLeuGlyHis	120
Db	577	ACAGTGGCTGCTCACGGCGCCCACTGCAGAAAGAAATTTCAGAGTCCGCTTCGGCCA	636
QY	120	sTYrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLYsSerIl	140
Db	637	CTACTCCCTGTACACAGTTATGTGATCTGGGCACACAGATGTTCACAGGGGGCAATCAT	696
QY	140	eProHisProLYrSerHisPrrGlnHisSerAsnAspLeuMetLeuIleLYsLeuAs	160
Db	697	CCCCCACCTGTGCTACTCCACCTGGCACCTTACACACTCATGTCTCATCAACTAA	756
QY	160	nArgArgIleArgPrrOThrLYsAspValArgProIlleAsnValSerSerHisCYsProSe	180
Db	757	CAGAAGAAATTGCTCCCACTAAAGATGTGAGACCATCAACGTCTCTCATTTGCTCCTC	816
QY	180	rAlaGlyThrLYsCYsIleuValSerGlyTyrGlyThrThrLYsSerProGlnValHisPrr	200
Db	817	TGCTGGGCAAAAGTCTGTGTGCTCGCGCTGGGGGCAACAAAGAGCCCCCAAGTGCAC	876
QY	200	eProLYsValIleuGlnCYsIleuAsnIlleSerValIleuSerGlnLYsArgLYsAspAl	220
Db	877	CCCTTAAGGTCTCCAGTCTGAATATACGGGTGTCTAAGTACGAAGAAAGTGGAGAGATGC	936
QY	220	aTYrProArgGlnIleAspAspThrMetPheCYsAlaGlyAspLYsAlaGlyArgAspSe	240
Db	937	TTACCCGAGACAGATPAGTACACCATGTTCTGGCGCGGTACAAAGCAGGTAGAGACTC	996
QY	240	rCYsGlnGlyAspSerGlyGlyProValValCYsAsnGlySerLeuGlnGlyLeuValSe	260
Db	997	CTGCCAGGGTATTTCTGGGGGGCGCTGTGGTCTGCATATGGCTCCCTGCAGGGACTCGTGT	1056
QY	260	rTYrGlyAspLYrProCYsAlaArgProAsnArgProGlyValTyrThrAsnLeuCYsLY	280
Db	1057	CTGGGGAAATTAACCTGTGTGCCCGGCCCAACAGACCGGGTGTCTTACAGAACTCTGCA	1116
QY	280	sPheThrLYrTrpIleGlnGlnThrIlleGlnAlaAsnSer 293	
Db	1117	GTTCACCAAGTGAATCCAGAAACATTCAGAGCCACTCC 1156	

```
; Sequence 1841, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ. ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1841
; LENGTH: 1143
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1841
Alignment Scores:
Pred. No.: 4,1e-150 Length: 1143
Score: 1479.00 Matches: 269
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 91.98% Indels: 0
DB: Gaps: 0
US-09-936-271b-14 (1-293) x US-09-949-016-1841 (1-1143)
QY 25 GlnH1eValLeuA1aAaNaSPVAlSeRCyAaPh1SPROSeRaenThrValProSeR 44
DB 6 GAGATGTTCTCGCCAAATGATGTTCTCGTGACACCCCTTAACAGCTGCTCT 65
QY 45 GlySerEnGlnApLeuGlyAlaGlyAlaGlyGluAaPAlaAaGSeRaSPaSPeSeR 64
DB 66 GGGGCAACACAGACCTCGGAGCTGGGGCCGGGGAACCCCGCTGGATGACAGAGC 125
QY 65 SerArg1le1leAnGlySeRaSPCyAaSPeR1sthrGlnProTPG1nAlaAlaLeu 84
DB 126 AGCGGCAATCAATGATGATCGAGTGGATGACACCCGCTGGAGCCGCTG 185
QY 85 LeuLeuArGPraEnGlnLeuTYCyG1YAlaValLeuValH1SPROG1nTPLeu 104
DB 186 TTGCTAAGGCCCAACAGCTCTACTGGGGGGGGTGTGGATCCACAGTGGCTGCTC 245
QY 105 ThrAlaAlaH1sCyAaRgLyAlaValaPheArGyAlaArgLeuGlyH1sTYRSeRLeuSeR 124
DB 246 ACCGCCCCCACTGACAGAAAGATTTCAGAGTCCGCTCGGCACTCTCCCTGCA 305
QY 125 ProValTYRGIuSeRGIyGInGlnMeRPhGInGlyVallySeR1leProH1aProGly 144
DB 306 CCACTTATGATGTGGGACAGAGATGTTCCAGGGGGGCAATCAATCCACCTGGGC 365
QY 145 TYSeRH1SPROG1YH1sSeRaSPaSPLeuMeRLeu1leYsLeuAaArGArG1leArG 164
DB 366 TACTCCCACTGGCCACTCTAAGACCTCACTGCTCAATCAAACTGAACAGAAATTGCT 425
QY 165 ProThrLySaPValaRPro1leAaSPaSPeSeRH1sCyAaSPROSeRa1AGlyThrLyS 184
DB 426 CCCCACTAAGATGTCAGACCCATCAAGCTCTCTCAATGTCTCTGCTGGGACAAAG 485
QY 185 CyLeuValSeRGIYTPG1YThrThrLySeSPROG1nValH1sPheProLyValLeu 204
DB 486 TGGTGGTGTGGCTGGGGACAAACAGAGCCCCCAAGTGCATTCCTTAAGGCTCTC 545
QY 205 GlnCySeuAaen1leSeRValLeuSeRGIuYsaRGIyAaPAlaTYRProArGIn 224
DB 546 CAGGCTTGATATCAAGGCTGCTAAGTCAGAAAGGTGCGAGATGCTTACCCAGACAG 605
```

```
QY 225 IleASPAPThrMeCPheCyAlaGlyAaPlyAlaGlyArGAPSeRCyGInGlyASP 244
DB 606 ATGATGACACCAATGTTCTGCGCGGTGACAAAGAGTGAAGATCTCTGCGAGGCTGAT 665
QY 245 SerGIYGIYProValaCyAaEnGlySeRLeuGInGlyLeuValSeRTripGIYArTYR 264
DB 666 TCGGGGGGGCTGTGGTGTGCAATGAGCTCCCTGACAGGAGACTGTGCTGGGGAGATTAC 725
QY 265 ProCYAlaArGPraEnArGPProGIYValTYRThraSPeR1leYsLeuCYAlaSPeRThrLySTP 284
DB 726 CCTTGTGCGCGGCGCCAAACAGACCGGGTGTCTTACACGAACCTTGCAAGTTCAACAAGTGG 785
QY 285 IleGlnGluThr1leGlnAlaAaSeR 293
DB 786 ATCCAGAAACATTCACAGGCCAATCTCC 812
RESULT 6
US-09-280-116-1
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ. ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-1
Alignment Scores:
Pred. No.: 1,03e-141 Length: 1504
Score: 1403.00 Matches: 280
Percent Similarity: 93.77% Conservative: 6
Best Local Similarity: 91.80% Mismatches: 7
Query Match: 87.25% Indels: 12
DB: Gaps: 2
US-09-936-271b-14 (1-293) x US-09-280-116-1 (1-1504)
QY 1 MetAlaThrAlaArGPProTPMeRTripValLeuCyAlaLeu1leThra1leLeu 20
DB 277 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTGTGCTGTGATCAACAGCTTGCTT 336
QY 21 LeuGlyValThrGlnH1sValLeuAlaAaNaSPVAlSeRCyAaPh1SPROSeRaSP 40
DB 337 CTGGGGGTCAACAGATGTTTCGCCAAACATGATGTTCTCGTGAACACCCCTCTAA 396
QY 40 nThrValProSeRGIYSeR--AaGlnApLeuGlyAla---GlyAlaGlyIuAaPAla 58
DB 397 CACCGTGGCTCTGGAGCAAAACAGACTTGGGAACCTGGGGCCCGGGGAAGAGGCC 456
QY 59 ArgSeRaSPaSPeSeRSeR--Arg1le1leAnGlySeRaSP--CyAaSPeR1sthrG 78
DB 457 CGGTGGATGACAGACAGACCGCGCATTCATATGATTCGAATCGGATGATACACACC 516
QY 78 InProTPG1nAla---AlaLeuLeuArGPraEnGlnLeu--TYRCyG1YAlaVala 96
DB 517 AGCGGTGGCAGGCGCGGTGTAAAGCCCAACCAAGTTTATTCGCGGGCGGT 576
QY 96 IleuValH1sPROG1nTPLeuLeuThrAlaAla-H1sCyAaRgLyAlaValaPheArGy 116
DB 577 GTTGGTGCATCCACAGAGGTGTTCAAGGGCCCACTGACAGAAAGATTTCAGAG 636
QY 116 AlaArgLeuGlyH1sTYRSeRLeuSeRProValTYRGIuSeRGIyGInGln-MetPhGIn 135
DB 637 TCGGTCTGGCCCAATTCCTCTGTCAACAGTTTATGATGTGGGACAGAAAGTTCAG 696
```

QY	136	ILlyVallysSerIleProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMet	155
Db	697	GGGGCAATCCATCCCTCCACCCTGGGCACTCCACCCTGGCCACTTAAAGACCTCATG	756
QY	156	LeuIleIysLeuAsnArgArgIleArgProThrIysAspValArgProIleAsnValSer	175
Db	757	CTCATCAAACTGAAACAGAAAGATTGTCCTCAATAAGATGTCAGACCCATTCAAGTCTCC	816
QY	176	SerHisCysProSerAlaGlyThrIlyCysLeuValSerGlyTyrPglYThrThrLysSer	195
Db	817	TCTCATTTCTCCCTGCTGGAGCAAAAGCTGGTGCTGGCTGGGGGACAAACAAGAC	876
QY	196	ProGlnValHisPheProIysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIys	215
Db	877	CCCCAAGGCACTTCCCTTAAGTCTCTCCAGTGCTTGAATATCAAGCGTCTTAAGTCAGAAA	936
QY	216	ArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIys	235
Db	937	AGTGTCCGAGAGTCTTACCCGACAGATGATGACACACTGTTCTGGCGCGGTGACAAA	996
QY	236	AlaGlnIysAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeu	255
Db	997	GAGGTAAAGACTCTGTCAGAGGATGATCTGGGGGCTGTGGTGTGCAATGAGCTCCCTG	1056
QY	256	GlnGlyLeuValSerTyrGlyAspTyrProCysAlaArgProAsnArgProGlyValTyr	275
Db	1057	CAGGAGCTCGTGTCCTGGGGGAGATTACCTTTGTGCCCGGCCCAACAGACCGGGGTGTCTA	1116
QY	275	rThAsnLeuIysIysPheThrIlyrTyrIleGlnGluThrIleGlnAlaAsnSer	293
Db	1117	CACGAACCTTCGCAAGTTCACCAAGTGGATCCAGAAAACATTCAGGCAACTCC	1171

RESULT 7

```

US-09-949-016-13583
; Sequence 13583, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13583
; LENGTH: 10818
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13583

```

Alignment Scores:

Pred. No.:	5.01e-97	Length:	10818
Score:	1002.00	Matches:	218
Percent Similarity:	44.13%	Conservative:	0
Best Local Similarity:	44.13%	Mismatches:	0
Query Match:	62.31%	Indels:	276
DB:	4	Gaps:	2

US-09-936-271B-14 (1-293) X US-09-949-016-13583 (1-10818)

Oy	25	GIuHISValLeuAlaAsnAsnAspValSerCysAspHisProSerAsnThrValProSer	44
Db	2006	GAGCATGTTCTCCGCGCMAATGATGTTTCTGTGACCAACCCCTTAACACCGTCCCTCT	2065
Oy	45	GlySerAsnGlnAspLeuGlyValAcGlyGlnAspAlaArgSerAspAspSerSer	64

Db	2066	GGAGACACCGACCTGGAGACTGGGGCCGGGAGAGCGCCGGTCGGATACAGCAGC	2126
Qy	65	SeArqIleIleasnGlySerAspCysaApwethIsthngInProTrpGlnAlaIleu	84
Db	2126	AGCGGCATCATCATATGATTCGACTGCGATATGCAACCCAGCCGTTGGCAGGCCCGCTG	2186
Qy	85	LeuLeuAtrProAsnGlnLeuTyrCysglYalaValIleuValHisProGlnITrPleuLeu	104
Db	2186	TTGGCTAAAGCCCAACCAAGCTCTACTCGCGGGCGGGTGTGGTCATCCACAGTGGCTGCTC	2244
Qy	105	ThraIaAlaHisCysArgLysLys-----	112
Db	2246	ACGGCCGCCACTGACGAGAGAGATGAAGTGGAGTTCCAAAGAGAGGGTTGGTGGGAGCG	2306
Qy	112	-----	112
Db	2306	GGGAAGTGGGGGTGGGGGTGGGGAAGTGGGGGTGGGGTGTGCATGAGAGTGAAGGCTGTG	2366
Qy	112	-----	112
Db	2366	GGGACGGGGAAGTGGGGTGGGGGTGCATGGAAGTGAAGGTTGTGGGATGGGTTG	2426
Qy	112	-----	112
Db	2426	GGGATGTGGAGCAGAGAGAGGTGCAGTTGGGGATAGGACTPAAGATGAGTTTTCGGG	2486
Qy	112	-----	112
Db	2486	GGAGCAAGSTGGAGAGATGAGGTTCGAGAGAGGAGAGTGTGTGATGGAAATGGGAAG	2546
Qy	112	-----	112
Db	2546	AGCCAAGATGGGTTCGATTTGGGGTTAGAGCATAATTTGTGAATGTTTGGGATGG	2606
Qy	112	-----	112
Db	2606	AGGTGGAATTGGATTGGCTTTAGAAATTGGGGGTGGGTAAGTAATCGGCTGGGCTGGA	2666
Qy	112	-----	112
Db	2666	TGAAGATAGCATGAGATAGGGTTGAGATTGGGAGCAGATATAGAAATGAAGATGGGAT	2726
Qy	112	-----	112
Db	2726	TGAGATTTGGGGTGGGGTTGAGATGTTGAGATTGGGCTTGAGAAATGCATATGGTGATG	2786
Qy	112	-----	112
Db	2786	GCTTCGGGTAGGGAAGAATTAGGGTTGGAAATGGATGGGTTTGAATTGTGACTGGG	2846
Qy	112	-----	112
Db	2846	ATGGGACAGCGCATGGGATTGGAACCAAGAGGAGTTGAGAGTGGTTTGGGACCGGGG	2906
Qy	112	-----	112
Db	2906	GTGGGGAATGGGGGTGGGGCTGGGGCTGGGTGCGGGTTGGGATGGCGTTGACGAGG	2966
Qy	113	-----ValPheArgValArgLe-----	118
Db	2966	ATTAGAGATCAGGGTTGGGTGACCTGGCCCATCTTCTCAGAGATTTTCAGAGTCCGCT	3026
Qy	118	uGlyHisIstYrSerIeuSerProValTyrGlnuSerGlyGlnGlnmetPheGlnGlyValY	138
Db	3026	CGGCGACACTACCTCGTGCACAGTTTAAGATCTGGGACGACGATGTTCCAGGGGGTCAA	3086
Qy	138	sSerIleProHisProGlyTyrSerHisProGlyHisSerAsnApLeuMetLeuIleuY	158
Db	3086	ATCATATCCCAACCTGGCTACTCCCAACCTGGCCACTCTPAACGACTCATGCTCATCA	3146
Qy	158	sIeuAsnArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCys	178


```
Db 250 CACAGCTTGTGAGGCCGACCAAGAGCCAGGAGCCAGATGTGTGAGGCCAGCTCTCCGTA 309
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspMetLeuIleValAsn 160
Db 310 CGGCACCCAGAGTACACAGACACCTTGTCTGCTACGACTTCTCATCAAGTTGGAC 369
Qy 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 370 GAATCCGTCGTCGACACACCATCCGAGCATGATGCTTCTGCGAGTCCCTTACC 429
Qy 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrIlySerProGlnValHisPhe 200
Db 430 GCGGGGAACCTCTTGGCTGCTTGTGTGGTGGGCTGCGCAACGCGCAAGAAG----- 483
Qy 201 ProIlyValIleuGlnCysLeuAsnIleSerValIleuSerGlnIlyAspGlyIleuAspAla 220
Db 484 CTTACCGTGTGAGTGTGACGCGGAAACGTCTCGTGTGTGTGAGAGGCTGTCAATTAAGCTC 543
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspIlyValArgAsp 239
Db 544 TATGACCCGCTGTACACACCCAGCATGTCTGTGCGCGGAGGAGGCAAGACCAAGAGAC 603
Qy 240 SerCysGlnIlyAspSerGlyIlyProValIlyCysAsnIlySerIleuGlnIlyLeuVal 259
Db 604 TCTTGAACGCTGACTCTGGGGGGCCCTGATCTGCAACGGGTACTTGCAGGGCCCTTGTG 663
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
Db 664 TCTTTCGGAAGAACCCCGTGTGGCCAAAGTTGGCGCTGCGCAAGTGTCTACACCAACTCTGC 723
Qy 280 LysPheThrIlyTrpIleGlnIlyThrIleGlnIlyAsn 292
Db 724 AATTCACGTAGATGATAGAGAAACCGTCCAGGCAAGT 762
```

RESULT 10

```
US-09-636-215-524.
; Sequence 524, Application US/09636215
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Reiter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-636-215-524
```

```
Alignment Scores:
Prod. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
```

```
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6
US-09-936-271B-14 (1-293) x US-09-636-215-524 (1-765)
```

```
Qy 1 MetAlaThrAlaArgProProTrpMetTrpValIleuCysAlaIleuIleThrAlaLeuLeu 20
Db 1 ARGCCACAGCAGGAAGAACCTCGGGGCTGTTCTCGGGGTACTCTATC----- 48
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 49 CTTGGTGTCT-----GCAGA 63
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyIlyGluAspAlaArgSer 60
Db 64 TCGCTGTCTGTGT----- 78
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 79 -----AGTCGACGCAATCATTAACGCGAGACTGACGCCGCACTCGACGCCCTGG 132
Qy 81 GlnAlaIleuLeuLeuArgProAsnGlnIleuTyrCysGlyAlaValIleuValHisPro 100
Db 133 CAGCGCGCACTGTGATGAA---AACGAATTGTTCTGCTCGGGGTCTGTCGATCCG 189
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgIlyValIleuValArgIleuGlnHis 120
Db 190 CAGTGGTGTCTGTACGCCGACACTGTTTCCAAACTCTTACACCACTCGGGCTGGCTTG 249
Qy 121 TyrSerLeuSerProValTyrIlySerGlyGlnIlyMetPheGlnIlyValIlySerIle 140
Db 250 CACAGCTTGAAGCCGACCAAGACGAGGAGCCAGATGTGTGAGGCCAGCTCTCCGTA 309
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAspAspMetLeuIleValAsn 160
Db 310 CGGCACCCAGAGTACACAGACCTTGTCTGCTGCAACCTCATCTCAAGTTGGAC 369
Qy 161 ArgArgIleArgProThrIlyAspValArgProIleAsnValSerSerHisCysProSer 180
Db 370 GAATCCGTCGTCGACACCATCCGAGCATGATGCTTCTGCGAGTCCCTTACC 429
Qy 181 AlaGlyThrIlyCysLeuValSerGlyTyrGlyThrIlySerProGlnValHisPhe 200
Db 430 GCGGGGAACCTCTGCTGCTTCTGTGGGCTGTGCGCAAGGCAAGT----- 483
Qy 201 ProIlyValIleuGlnCysLeuAsnIleSerValIleuSerGlnIlyAspGlyIleuAspAla 220
Db 484 CTTACCGTGTGACGTCGTAACGTGTGGGTGTGTGAGAGAGTCTGCAATTAAGCTC 543
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspIlyValArgAsp 239
Db 544 TATGACCCGCTGTACACCCAGCATGTCTGCGCGGAGGAGGCAAGACCAAGAGAC 603
Qy 240 SerCysGlnIlyAspSerGlyIlyProValIlyCysAsnIlySerIleuGlnIlyLeuVal 259
Db 604 TCTTGAACGCTGACTCTGGGGGGCCCTGATCTGCAACGGGTACTTGCAGGGCCCTTGTG 663
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValIlyThrAsnLeuCys 279
Db 664 TCTTTCGGAAGAACCCCGTGTGGCCAAAGTTGGCGCTGCGCAAGTGTCTACACCAACTCTGC 723
Qy 280 LysPheThrIlyTrpIleGlnIlyThrIleGlnIlyAsn 292
Db 724 AATTCACGTAGATGATAGAGAAACCGTCCAGGCAAGT 762
```

RESULT 11

```
US-09-685-166A-524
; Sequence 524, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
```

```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-685-166A-524

Alignment Scores:
Pred. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
Gaps: 6

US-09-936-271B-14 (1-293) x US-09-685-166A-524 (1-765)
QY 1 MetAlaThrAlaArgProPOTrMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
Db 1 ATGGCCACAGCAGGAAATCCCTGGGCTGTTCTGGGCTGTAACCTC----- 48
QY 21 LeuGlyValThrGluIleValLeuAlaAsnAspValSerCyAspAspHisProSerAsn 40
Db 49 CTGGGTCTC-----GCGGGA 63
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAlaAspValArgSer 60
Db 64 TCGCTGCTCTGCT----- 78
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspAspMetHisThrGlnProTrp 80
Db 79 -----AGCTGCAGCCAAATCATTAACGGCGAGGACTGCGAGCCCGCAGCTCGCAGCCCTGG 132
QY 81 GlnAlaIleLeuLeuLeuAspArgProAsnGlnLeuTrpCyAsnGlyAlaValLeuValHisPro 100
Db 133 CAGCGGCACTGGTCATGAA--AACGAATTGTTCTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTG 189
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspArgValPheAspValArgLeuGlyHis 120
Db 190 CAGTGGGTGGTGTGAGCGGCACTGTTTCCAGAACTCTTACACATCGGCGCTGGGCTG 249
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValysSerIle 140
Db 250 CACAGTCTTGAAGCGGCAACAAGCCAGGAGCGGAGTGGTGGAGGCGCAGCTTCCTGTA 309
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValLeuAsn 160
Db 310 CGGCAACCCAGAGTCAACAGACCCCTGCTGCTTAAAGCACTCAAGCTCATTAAGTTGAC 369
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyAspProSer 180
Db 370 GAATCCGATGCGAGTGTGACACATCCGAGCATTCAGCATTTGCTTGGCAGTGCCTTACC 429
```

```

QY 181 AlaGlyThrIleCyAspLeuValSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
Db 430 GCGGGAACTTGTCTCTGTTTGGCTGGGCTCTCTGCGCAACGCGAAGT----- 463
QY 201 ProLeuValLeuGlnCyAspLeuHisSerValLeuSerGlnIleAspArgCyAsnValAsp 220
Db 484 CTTACCGTGTGTGAGTGGCGGAAAGTGTGCTGTGTGTGAGAGGTCTGAGTAACTGAC 543
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAsnLeuGlyAspValArgIleAsp 239
Db 544 TATGACCGGCTGTACCAACCCAGCATGTTCTGGCGCGGAGCGAAGACGCAAGAGAC 603
QY 240 SerCyGlnGlnIleAspSerGlyIleProValIleCyAsnGlySerLeuGlnIleVal 259
Db 604 TCTTGCAACGCTACTGTGGGGGGCCCTGATGTGCAACGGAATTCGACGGGCTTGTG 663
QY 260 SerTrpGlyAspTyrProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCy 279
Db 664 TCTTGGAAAGCCCGGTGTGGCCAAAGTTGGCGTGCAGGTGTCTTACACCACTCTGC 723
QY 280 LysPheThrIleTrpIleGlnGlyIleGlnAlaAsn 292
Db 724 AAATTCATGATGATAGAAAGAACGTCACAGCAGT 762

RESULT 12
US-09-679-426-524
; Sequence 524, Application US/09679426
; Patent No. 6759515
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqun
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-679-426-524

Alignment Scores:
Pred. No.: 5,26e-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
Gaps: 6

US-09-936-271B-14 (1-293) x US-09-679-426-524 (1-765)
QY 1 MetAlaThrAlaArgProPOTrMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
Db 1 ATGGCCACAGCAGGAAATCCCTGGGCTGTTCTGGGCTGTAACCTC----- 48
```

```
QY 21 LeuGIvValThrGIunhIsvaLleuAlaAsnAspValSerCyAspHisProSerAn 40
Db 49 CTGGGTCTC-----GCAGCA 63
QY 41 ThrValProSerGIySerAenGlnAspLeuGIyAlaGIyAlaGIyGlnAspAlaArgSer 60
Db 64 TCCTGCTCTCTCTG----- 78
QY 61 AspAspSerSerSerArgIleIleAsnGIySerAAspCyAspMetHisThrGlnProTrp 80
Db 79 -----AGCTGCAGCCAAATCATAAACGGGAGAGACTGCAGCCCGCAGCTGCAGCCCTGG 132
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCyAsnGIyAlaValLeuValHisPro 100
Db 133 CAGCGCGCACTGCTCATGAA---AACGAATTCTTCTGCTCGGCGCTCTCTGCTCATCCG 189
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspArgLyAspValPheArgValArgLeuGIyHis 120
Db 190 CAGTGGGTGCTGTCAAGCCGACACTGTTTCAGAACTCTTAACACATCGGGCTGGGCTGG 249
QY 121 TyrSerLeuSerProValTyrGIySerGIyGlnGlnMetPheGlnGlnValYValysSerIle 140
Db 250 CACAGCTTGAAGCCGACCAAGAGCCAGAGGCAAGATGGTGAAGGCCAGCCTCTCTCGTA 309
QY 141 ProHisProGIyTyrSerHisProGIyHisSerAspAspLeuMetLeuIleLyAsn 160
Db 310 CGGCACCCAGAGTAAACAAGACACCTTGTCTGTCTCAACACTCTCTCATCAAGTTGGAC 369
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyAspProSer 180
Db 370 GAATCCGTGCCAGTCTGACACCATCCGAGCATGACATGCTTCCGAGTGCCTTACC 429
QY 181 AlaGIyThrLySerLyLeuValSerGIyTrpGIyThrThrLySerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGCTCTGTTTGTGGTGGGCTGCTGGCGAAGCGCAGAAAG----- 483
QY 201 ProLyValLeuGlnCyLeuAsnIleSerValLeuSerGlnTyrAspCyAsnValAspAla 220
Db 484 CTTACCGTGTGAGTGAAGTGGTGAACGTGTGCTGTCTGAGAGGTCTCTCACTTAAGCTC 543
QY 221 TyrProArgGlnIleAspAspThrMetPheCyAsnIleGlyAspValArgIleArgAsp 239
Db 544 TATGACCCCGCTGTACACCCACCATGTTCTGCGCGCGAGGAGCAAGACAGAGAGAC 603
QY 240 SerCyGlnGlnLyAspSerGIyGIyProValValCyAsnGIySerLeuGlnGIyVal 259
Db 604 TCTTGCAACGGTACTCTGGGGGAGCCCTGATCTGCAACGGGTACTTGCAAGGCTTGTG 663
QY 260 SerTrpGIyAspTyrProCyAsnIleArgProAsnArgProGIyValTyrThrAsnLeuCy 279
Db 664 TCTTGGAAAAAGCCCGTGTGGCCAAAGTTGGGCTGCAGGTCTCTACCAACCTCTGCT 723
QY 280 LyAspPheThrLySerTrpIleGlnGlnIleValHisPhe 292
Db 724 AAATTCATGAGTGATAGAGAAACCGTCCAGGCGAGT 762

RESULT 13
US-09-759-143-524
Sequence 524, Application US/09759143
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yuxui
APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
```

```
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel
APPLICANT: Mang, Aijun
APPLICANT: Skeiky, Yashir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C23
CURRENT APPLICATION NUMBER: US/09/759,143
NUMBER OF SEQ ID NOS: 934
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 524
LENGTH: 765
TYPE: DNA
ORGANISM: Homo sapien
US-09-759-143-524

Alignment Scores:
Pred. No.: 5,266-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6

US-09-936-271B-14 (1-293) x US-09-759-143-524 (1-765)
QY 1 MetaIaThrAlaArgProTrpMetTrpValLeuCyValaLeuIleThrAlaLeuLeu 20
Db 1 ATGGCACAAGCAGAAATCCCTGGGGCTGGTCTCGGGATACCTATC----- 48
QY 21 LeuGIvValThrGIunhIsvaLleuAlaAsnAspValSerCyAspHisProSerAn 40
Db 49 CTGGGTCTC-----GCAGCA 63
QY 41 ThrValProSerGIySerAenGlnAspLeuGIyAlaGIyAlaGIyGlnAspAlaArgSer 60
Db 64 TCCTGCTCTCTCTG----- 78
QY 61 AspAspSerSerSerArgIleIleAsnGIySerAAspCyAspMetHisThrGlnProTrp 80
Db 79 -----AGCTGCAGCCAAATCATAAACGGGAGAGACTGCAGCCCGCAGCTGCAGCCCTGG 132
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCyAsnGIyAlaValLeuValHisPro 100
Db 133 CAGCGCGCACTGCTCATGAA---AACGAATTCTTCTGCTCGGCGCTCTCTGCTCATCCG 189
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspArgLyAspValPheArgValArgLeuGIyHis 120
Db 190 CAGTGGGTGCTGTCAAGCCGACACTGTTTCAGAACTCTTAACACATCGGGCTGGGCTGG 249
QY 121 TyrSerLeuSerProValTyrGIySerGIyGlnGlnMetPheGlnGlnValYValysSerIle 140
Db 250 CACAGCTTGAAGCCGACCAAGAGCCAGAGGCAAGATGGTGAAGGCCAGCCTCTCTCGTA 309
QY 141 ProHisProGIyTyrSerHisProGIyHisSerAspAspLeuMetLeuIleLyAsn 160
Db 310 CGGCACCCAGAGTAAACAAGACACCTTGTCTGTCTCAACACTCTCTCATCAAGTTGGAC 369
QY 161 ArgArgIleArgProThrTyrAspValArgProIleAsnValSerSerHisCyAspProSer 180
Db 370 GAATCCGTGCCAGTCTGACACCATCCGAGCATGACATGCTTCCGAGTGCCTTACC 429
QY 181 AlaGIyThrLySerLyLeuValSerGIyTrpGIyThrThrLySerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGCTCTGTTTGTGGTGGGCTGCTGGCGAAGCGCAGAAAG----- 483
QY 201 ProLyValLeuGlnCyLeuAsnIleSerValLeuSerGlnIleValArgCyAsnValAspAla 220
Db 484 CTTACCGTGTGAGTGAAGTGGTGAACGTGTGCTGTGAGAGGTCTCTCACTTAAGCTC 543
```


Qy 221 TyrProArgGlnIleAspThrMetPheCysAla---GlyAspValAlaGlyArgAsp 239
Db 544 TATACCCCGCTGTACACCCAGCATGTTCTGCGCGGAGGAGGACCAAGAC 603
Qy 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
Db 604 TCCTGCACACGGTGACTCTGGGGGGCCCTGATCTGCACACGGGTACTTGCACAGGCTTGTG 663
Qy 260 SerTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
Db 664 TCCTTCGAAAGAGCCCGCTGTGGCCAGTTGCGCGGTGCTCTACCAACCTCTGCG 723
Qy 280 LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn 292
Db 724 AAATTCACTGATGATAGAGAAACCGTCCAGGCGCACT 762

RESULT 14
US-09-651-236-524
Sequence 524, Application US/09651236
Patent No. 6818751
GENERAL INFORMATION:
APPLICANT: Xu, Jiaangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitchell, Jennifer L.
APPLICANT: Heflocker, Susan L.
APPLICANT: Jiang, Yuqun
APPLICANT: Henderson, Robert A.
APPLICANT: Kadios, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Rether, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Derrick
APPLICANT: Li, Samuel
APPLICANT: Wang, Aijun
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121-42718C18
CURRENT APPLICATION NUMBER: US/09/651,236
CURRENT FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 865
SOFTWARE: PatSeq for Windows Version 3.0
SEQ ID NO 524
LENGTH: 765
TYPE: DNA
ORGANISM: Homo sapien
US-09-651-236-524

Alignment Scores:
Pred. No.: 5,266-63 Length: 765
Score: 671.00 Matches: 134
Percent Similarity: 61.77% Conservative: 47
Best Local Similarity: 45.73% Mismatches: 72
Query Match: 41.73% Indels: 40
DB: 4 Gaps: 6

US-09-936-271b-14 (1-293) x US-09-651-236-524 (1-765)

Qy 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
Db 1 ATGGCCACAGAGAAATCCCTGGGCGGTCTTCTGGGTACTCTATC-----48
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn 40
Db 49 CTGTGTGTC-----GCAAGA 63
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAspAlaArgSer 60
Db 64 TCCTCGCTCTCTGCT-----78

Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 79 -----AGCTGCAGCCAAATCATTAAGGAGAGACTGCGCCGCACTGCAAGCTTGG 132
Qy 81 GlnAlaIleLeuLeuAspArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 133 CAGCGCGCACTGGTCATGAAA---AACGAATTGTTCGTCTGGGCGGTCTGGTGCATCCG 169
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgValLeuValPheArgValAlaGlnGlyHis 120
Db 190 CAGTGGGTGTCTGACCGCACCTGTTCCTCAAACTCCACACCACTGGGCTGGGCTG 249
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 250 CACAGTTTAAAGGCCAGCAAGAGCCAGGAGCCAGATGTGAGAGCCACCTCTCCGTA 309
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 310 CGGACCCACAGTACAAACAGACCTTGTCTGCTAGACCTCATGTCTCATAGTTGGAC 369
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 370 GAATCCGTGCCAGTCTGACACCATCCGAGCATCAGCATGCTTCCGACGTCCCTAC 429
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyTyrThrLysSerProGlnValHisPhe 200
Db 430 GCGGGAACTCTTGCTCTGCTGGGCTGCTGCGGAGCAAGCCAGAAAG-----483
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 484 CTTACCGTCTGCAGTCCGGAACGTCTCGTGTGTCTGAGAGGTCTCAGTAAGCTC 543
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspValAlaGlyArgAsp 239
Db 544 TATGACCCCGCTGTACACCCAGCATGTTCTGGCGCGGCGGAGGAGGACCAAGAC 603
Qy 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
Db 604 TCCTGCACAGTGACTCTGGGGGCCCTGATCTGCACAGGATCTTGCAGGGCTTGTG 663
Qy 260 SerTrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279
Db 664 TCCTTCGAAAGAGCCCGTGTGGCCAGTTGCGCGGTGCTCTACCAACCTCTGCG 723
Qy 280 LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn 292
Db 724 AAATTCACTGATGATAGAGAAACCGTCCAGGCGCACT 762

RESULT 15
US-09-386-642-10
Sequence 10, Application US/09386642
Patent No. 6420157
GENERAL INFORMATION:
APPLICANT: Darrow, Andrew
APPLICANT: Qi, Jenson
APPLICANT: Andrade-Gordon, Patricia
TITLE OF INVENTION: Zymogen Activation System
FILE REFERENCE: ORT-1028
CURRENT APPLICATION NUMBER: US/09/386,642
CURRENT FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 1052
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-10

Alignment Scores:
Pred. No.: 5,236-61 Length: 1052

Score: 654.50 Matches: 129
 Percent Similarity: 62.99% Conservative: 48
 Best Local Similarity: 45.91% Mismatches: 87
 Query Match: 40.70% Indels: 17
 DB: 3 Gaps: 7

US-09-936-271b-14 (1-293) x US-09-386-642-10 (1-1052)

```

QY      19  LeuLeuLeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisPro 38
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      49  CTCCTGCTGCTGGTGTGTCCTAATCTACTTGTGCCAGGGTGTGCTCTCCGACTACAG 108
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      39  SerAsnThrValProSerGlySerAsnGlnAspLeuGlyValGlyValGlyGluAspAla 58
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      109 -----GACGACGACGACGTCGAGCGCGCT-----CTTGCT 141
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      59  ArgSerAspAspSerSerSerArgLleIleAsnGlySerAspCys--AspMetHisThr 77
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      142  GCCCCTTGTGATGATGATGACAAAGATCGTTGGGGGCTACAACTGTCTAGAAAGCACTCC 201
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      78  GlnProTrpGlnAlaAlaLeuLeuLeuArgProAsnGlnLeuThrCysGlyValValLeu 97
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      202  CAGCCCTCGGAGGAGCC--CTGTCGAGAGACGGGCTACTCTGTGGGGGAGCGCTC 258
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      98  ValHisProGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArg 117
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      259  ATGCCCCGAGATGGCTCTGACAGAGCCCACTGCTCAAGCCCGCTACATAGTTAC 318
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      118  LeuGlyHisTrpSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyVal 137
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      319  CTGGGGACAGCAACCTCCAGAG--GAGAGGGCTGTGACAGACCCGACAGCCACT 375
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      138  LysSerLleProHisProGlyTyrSer-----HisProGlyHisSerAsnAsp 153
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      376  GAGTCTCTCCCAACCCCGCTTCAACAAGACCTCCCAACAAGACACCGCAATGAC 435
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      154  LeuMetLeuLleLysLeuAsnArgArgLleArgProThrLysAspValArgProIleAsn 173
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      436  ATCATGCTGTGAAGATGGCATGCCAGTCTCATCATCAGGAGCTGTGGAGCCCTCAC 495
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      174  ValSerSerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThr 193
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      496  CTCCTCTCAGCTGTGTGCTGCTGCGACCACTGCTCATTTCCGCTGGGGCAGCAGC 555
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      194  LysSerProGlnValHisPheProLysValLeuGlnCysLeuAsnLleSerValLeuSer 213
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      556  TCCAGCCCCCAGTTACGCTGCTCACACCTTGAGTGCAGCCCAACATCACCATCATTTGAG 615
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      214  GlnLysArgCysGlnAspAlaTyrProArgGlnLleAspAspThrMetPheCysAlaGly 233
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      616  CACCAGAAAGTGAAGACGCTACCCCGGCAACATCACAGACATGATGTTGCCAGC 675
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      234  ---AspLysAlaGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsn 252
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      676  GTCGACGAGAGGGGCAAGACTCTGCGACAGTGACTCCCGGGGCCCTGTGTCTGTAAAC 735
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      253  GlySerLeuGlnGlyLeuValSerTrpGlyAspTyrProCysAlaArgProAsnArgPro 272
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      736  CAGTCTCTTCAAGGCATTAATCTCTGGGGCCAGAGATCCGTGTGGATCACCCGAAAGCT 795
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      273  GlyValTyrThrAsnLeuCysLysPheThrLysTrpLleGlnGlnLutThrIleGlnAlaAsn 292
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      796  GGTGTCTACACGAAGCTCAATATATGTGACTGGATTCAGAGAGCATGAAGAACAAAT 855
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      293  Ser 293
        |||
DB      856  TCT 858
        |||

```

Search completed: February 26, 2005, 21:27:34
 Job time : 211 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:48:31 : Search time 634 Seconds

(without alignments)
2738.422 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608
Sequence: 1 MATRRPMMVLCALITALL.....VTNNLCKTKMTQETIQANS 293

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q=/cg2_1/USPTO.epool/US09936271/runat_23022005_130128_27307/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rmb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRI=blosum62
-TRANS=human40.csl -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPEIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09936271.0CGN_1_1480 @runat_23022005_130128_27307
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cg2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cg2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cg2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
18: /cg2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
19: /cg2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cg2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cg2_6/prodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cg2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1608	100.0	1260	17	US-10-172-118-1411
2	1608	100.0	1260	17	US-10-295-027-515
3	1608	100.0	1260	17	US-10-173-999-53
4	1608	100.0	1260	17	US-10-342-887-1411
5	1608	100.0	1381	9	US-09-739-907-52
6	1608	100.0	1381	11	US-09-938-671-52
7	1608	100.0	1516	9	US-09-739-907-37
8	1608	100.0	1516	11	US-09-938-671-37
9	1608	100.0	1570	9	US-09-989-722-308
10	1608	100.0	1570	9	US-09-989-723-308
11	1608	100.0	1570	9	US-09-989-729-308
12	1608	100.0	1570	9	US-09-989-727-308
13	1608	100.0	1570	9	US-09-989-731-308
14	1608	100.0	1570	9	US-09-989-732-308
15	1608	100.0	1570	9	US-09-991-073-308
16	1608	100.0	1570	9	US-09-990-442-308
17	1608	100.0	1570	9	US-09-991-163-308
18	1608	100.0	1570	9	US-09-993-604-308
19	1608	100.0	1570	9	US-09-990-456-308
20	1608	100.0	1570	9	US-09-989-721-308
21	1608	100.0	1570	9	US-09-992-598-308
22	1608	100.0	1570	9	US-09-989-293-308
23	1608	100.0	1570	9	US-09-989-735-308
24	1608	100.0	1570	9	US-09-990-444-308
25	1608	100.0	1570	9	US-09-991-181-308
26	1608	100.0	1570	9	US-09-989-730-308
27	1608	100.0	1570	9	US-09-990-436-308
28	1608	100.0	1570	9	US-09-993-687-308
29	1608	100.0	1570	10	US-09-989-724-308
30	1608	100.0	1570	10	US-09-997-653-308
31	1608	100.0	1570	10	US-09-989-724-308
32	1608	100.0	1570	10	US-09-989-728-308
33	1608	100.0	1570	10	US-09-990-441-308
34	1608	100.0	1570	10	US-09-993-667-308
35	1608	100.0	1570	10	US-09-997-428-308
36	1608	100.0	1570	10	US-09-997-666-308
37	1608	100.0	1570	10	US-09-990-438-308
38	1608	100.0	1570	10	US-09-990-452-308
39	1608	100.0	1570	10	US-09-990-711-308
40	1608	100.0	1570	10	US-09-989-726-308
41	1608	100.0	1570	10	US-09-989-156-308
42	1608	100.0	1570	10	US-09-990-437-308
43	1608	100.0	1570	10	US-09-991-157-308
44	1608	100.0	1570	10	US-09-997-514-308
45	1608	100.0	1570	10	US-09-997-573-308

ALIGNMENTS

RESULT 1
US-10-172-118-1411
; Sequence 1411, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699

```
; SEQ ID NO 1411
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_012427
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-1411
```

Alignment Scores:

Pred. No.:	3,29e-187	Length:	1260
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-09-936-271B-14 (1-293) x US-10-172-118-1411 (1-1260)

```
QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
DB 43 ATGGCTACAGCAAGACCCCTGGATGTGGTCTCTGCTCTGATCAGCCTTGCTT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSer 40
DB 103 CTGGGGGTACAGAGCATGTTCTGCCCAACATGATGTTCTCTGACCACTTAAC 162
QY 41 ThrValProSerGlySerSerGlnAspLeuGlyValGlyValGluAspSer 60
DB 163 ACCGTCCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGGAAGACGCCCTG 222
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnPro 80
DB 223 GATGACAGCAGCAGCGGCATCATCATGATCCGATCGGATATGACACCCAGCCTG 282
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCyGlyValAlaValLeuValHisPro 100
DB 283 CAGGCGCGCTGTGCTAAGGCCCAACAGCTTACTGGGGGGCTGTGGTGCATCA 342
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyValValPheArgValArgLeuGlyHis 120
DB 343 CAGTGGCTGTCAAGCGCGCCCACTCAAGAAAGATTTCAGATCCGCTCGGCCAC 402
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
DB 403 TACTCCCTGTCAACAGTTATGATATGAGGAGCAATGTTCCAGGGGGTCAATCATC 462
QY 141 ProHisProGlyTyrTyrSerHisProGlyHisSerAsnAspLeuLeuIleValLeuAsn 160
DB 463 CCCCACTGGGTACTCCACCTCCGCACTTAAAGACCTCATGCTCAAACTGAAC 522
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 523 AAGAGATTGTCTCCCTAAAGATGTCAGACCCATCAAGCTCTCTCAATTGCTCTCT 582
QY 181 AlaGlyThrIleCysLeuValSerGlyTyrGlyThrThrIleSerProGlnValHisPhe 200
DB 583 GCGGGGCAAAAGTGTGATGTCTGGCTGGGGGCAACCAAGAGCCCCAAAGTCACTTC 642
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGluAspAla 220
DB 643 CCTAAGGTCTCCAGTGTGAATATCAAGCTGCTAAGTCAGAAAAGTGCAGATGCT 702
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysValIleGlyArgAspSer 240
DB 703 TACCCAGACAGATATGATCAACCATGTCTGGCGGGGACAAACAGGTGAGACTCC 762
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValIleValSer 260
DB 763 TGCAGGGGTGATTTGGGGGGCTGTGTGTGCAAGGCTCCCTGAGAGGACTCGGTGC 822
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
```

```
DB 823 TGGGAGATTAACCTTGTGCCCGCCCAACAGACCGGGTGTCTACAGAACCTTGCAAG 882
QY 281 PheThrLysTrpIleGlnGluThrIleGlnIleAsnSer 293
DB 883 TTCACCAAGTGTATCCAGGAACATCCAGGCCAACTCC 921
```

RESULT 2

US-10-295-027-515

Sequence 515, Application US/10295027

Publication No. US20030232350A1

GENERAL INFORMATION:

```
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natacha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezil, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; PRIORITY FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIORITY FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 515
```

LENGTH: 1260

TYPE: DNA

ORGANISM: Homo sapiens

US-10-295-027-515

Alignment Scores:

Pred. No.:	3,29e-187	Length:	1260
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	17	Gaps:	0

US-09-936-271B-14 (1-293) x US-10-295-027-515 (1-1260)

```
QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCyAlaLeuIleThrAlaLeu 20
DB 43 ATGGCTACAGCAAGACCCCTGGATGTGGTCTCTGCTCTGATCAGCCTTGCTT 102
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSer 40
DB 103 CTGGGGGTACAGAGCATGTTCTGCCCAACATGATGTTCTCTGACCACTTAAC 162
```



```
Db      823 TGGGAGATTACTTGTGCCCCGCCAACAGACCGGGTGTCTACAGAACTCTTGCAAG 882
Qy      281 PheThrIyETripIeGIngluThriIegiIaIaenSer 293
Db      883 TTCACCAAGTGGATCCAGAAACCATCCAGGCCACTCC 921

RESULT 4
US-10-342-887-1411
/ Sequence 1411, Application US/10342887
/ Publication No. US20040058340A1
/ GENERAL INFORMATION:
/ APPLICANT: Dai, Hongyue
/ APPLICANT: He, Yudong
/ APPLICANT: Linsley, Peter S.
/ APPLICANT: Mao, Mao
/ APPLICANT: Roberts, Christopher J.
/ APPLICANT: Van 't Veer, Laura Johanna
/ APPLICANT: Van de Vijver, Marc J.
/ TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
/ FILE REFERENCE: 9301-188-999
/ CURRENT APPLICATION NUMBER: US/10/342,887
/ CURRENT FILING DATE: 2003-01-15
/ PRIOR APPLICATION NUMBER: 60/298,918
/ PRIOR FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: 60/380,710
/ PRIOR FILING DATE: 2002-05-14
/ PRIOR APPLICATION NUMBER: 10/172,118
/ PRIOR FILING DATE: 2002-06-14
/ NUMBER OF SEQ ID NOS: 2699
/ SEQ ID NO 1411
/ LENGTH: 1260
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-342-887-1411

Alignment Scores:
Pred. No.:      3,29e-187      Length:      1260
Score:          1608.00      Matches:      293
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              17          Gaps:          0

US-09-936-271b-14 (1-293) x US-10-342-887-1411 (1-1260)
Qy      1 MetAlaIaThAlaArGProTrpMetTTPvAlleuCyAlaIaIeulIeThAlaIeuleu 20
Db      43 ATGGCTACAGCAAGACCCCTGGAGTGGGTCTGTGCTGTATCAACACCTTGCTT 102
Qy      21 LeuGlyValThrGlUhiSvAlIeulIaAaAnaAspValSerCyAspHisProSerAan 40
Db      103 CTGGGGGTCAAGAGATGTTCTCGCCACATGATGTTTCTGTACACCCCTTAAAC 162
Qy      41 ThrValProSerGlySerAnGlnAaPleuGlyAlaGlyValGlyLnuAspAlaArGSer 60
Db      163 ACCGGCTCCTGGGAGCAACCAAGACCTGGAGCTGGGGCGGGAGAACGCCGCTGCG 222
Qy      61 AspAspSerSerSerArgIleIleAnGlySerAspCyAspMetHisThrGlnProTrp 80
Db      223 GATGACAGACGACCGCATCATCATGATCCGACTGGCAATATGACACCCAGCCGCTGG 282
Qy      81 GlnAlaAlaIeuleuArGProAnGlnIeulIeUyTyCyGlyValaValIeulValHisPro 100
Db      283 CAGGCCGCTGTGTTAAGGCCCAACAGCTTACTTGGGGGGGTGTGGTGCATCCA 342
Qy      101 GlnTrpLeuIeulThrAlaAlaHisCyArGlySlyValPheArGValArGleuGlyHis 120
Db      343 CAGTGGCTGCTACAGCCGCCCACTGCAGAGAAAGATTTCAGATCGTCTCGGCCAC 402
Qy      121 TyrSerIeulSerProValTyrgIuSerGlyGlnGlnIeulMetPheGlnGlyValIysSerIle 140
Db      403 TACTCCCTGTCAACAGTTATGATCTGGGACAGCATGTTCCAGGGGGTCAATCCATC 462
```

```
Qy      141 ProHisProGlyTySerHisProGlyHisSerAAnaPleuMetIeulIeLysLeuAn 160
Db      463 CCCACCTGGGTACTCCACACCTGGCCACTTAACGACCTCATCTCATCAACTGAAAC 522
Qy      161 ArgArgIleArgProThrIyAspValArgProIleAnaValSerSerHisCyAsProSer 180
Db      523 AGAAGAATTGCTCCACTTAAGATGTACAGACCCATCAAGTCTCTCATTTGCTCTCT 582
Qy      181 AlaGlyThrIyCyLeuValSerGlyTrpGlyThrIyHisSerProGlyValHisPhe 200
Db      583 GCTGGACAAAGTGTGTGTGTGCTGGGGGACACCAAGAGCCCAAGTGACTTC 642
Qy      201 ProLysValIeulGlnCyLeuAnIleSerValIeulSerGlnIyArGlyAspGlyAspAla 220
Db      643 CCTAAGTCTCTCACTGCTTGAATTCAGCGTCAATGCAAAAAGGTGCGAGGATGT 702
Qy      221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspLysAlaGlyArgAspSer 240
Db      703 TACCCGACAGATGATGACACCATGTTGCGCGGTGACAAAGCAGGTAGAGACTCC 762
Qy      241 CySGInGlyAspSerGlyGlyProValValCyAsnGlySerIeulGlnIyLeuValSer 260
Db      763 TCCAGGGGTGATCTGGGGGGCTGTGTGTGCAATGTCTCTCCAGAGGACTGTGTCTC 822
Qy      261 TrpGlyAspTyProCyAlaArGProAnaArGProGlyValTyTrpAsnLeuCySerIys 280
Db      823 TGGGAGATTACCTCTGTGCCCCGCCAACAGACCGGGTGTCTACACAACTCTGCAAG 882
Qy      281 PheThrIyETripIeGIngluThriIegiIaIaenSer 293
Db      883 TTCACCAAGTGGATCCAGAAACCATCCAGGCCACTCC 921
```

```
RESULT 5
US-09-739-907-52
/ Sequence 52, Application US/09739907
/ Patent No. US20010012889A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 36 Human Secreted Proteins
/ FILE REFERENCE: P2022P1
/ CURRENT APPLICATION NUMBER: US/09/739,907
/ CURRENT FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 09/348,457
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/070,567
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070,692
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070,704
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070,658
/ PRIOR FILING DATE: 1998-01-07
/ NUMBER OF SEQ ID NOS: 196
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 52
/ LENGTH: 1381
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-739-907-52

Alignment Scores:
Pred. No.:      3,73e-187      Length:      1381
Score:          1608.00      Matches:      293
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              9          Gaps:          0

US-09-936-271b-14 (1-293) x US-09-739-907-52 (1-1381)
Qy      1 MetAlaThrAlaArGProTrpMetTTPvAlleuCyAlaIaIeulIeThAlaIeuleu 20
```

```

Db      129 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTGCTGTGATCACAGCTTCTT 188
QY      21 LeuGIYValThrGIuHIEValLeuAlaAspAsnAspValSerCyAspAspHISProSerAsn 40
      189 CTGGGGGTGCACAGAGCATGTCTTCGCGCAACAATATGTTCTCTGTGACCAACCCCTCTTAC 248
QY      41 ThrValProSerGIYSerAsnGlnAspLeuGIYValaGIYValaGIYValaAspAlaArgSer 60
      249 ACCGTGCTCTTGGAGACAAACAGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 308
Db      61 AspAspSerSerSerArgIleIleAsnGIYSerAspCyAspMetHisThrGlnProTrp 80
      309 GATGACAGACAGACCGCATCATCATATGATTCGACATCCGATATGACACCCAGCCCTGG 368
QY      81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTrpCyseGIYValaValLeuValHisPro 100
      369 CAGAGCCCGCTGTGTGCTTAAAGCCCAACAGCTTACGCGGGGGGTGTGTGATCACTCA 428
QY      101 GlnTrpLeuLeuThraAlaAlaHisCyAspArgLysValPheArgValArgLeuGIYHis 120
      429 CAGTGGCTGCTCACGGCGGCGCCACTGCAAGAAAGTTTTCAGAGTCCGTTCTGGGCAC 488
QY      121 TyrSerLeuSerProValTYRGIUSerGIYGIInIleMetPheGlnGIYValIYsSerIle 140
      489 TACTCCCTGTCCACAGTTTATGATCTGGGACAGATGTTTCCAGGGGGTCAAAATCATC 548
Db      141 ProHisProGIYTYRSerHisProGIYHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
      549 CCCACCCCTGCTACTCCACCCCTGGCCACTTAAACCTCATGCTCATCAAACTGAC 608
QY      161 ArgArgIleArgProThraLysAspValArgProIleAsnValSerSerHisCyseProSer 180
      609 AGAAGATTGTGTCCACTTAAAGATGACAGCCCAACAGCTCTCTCATGTGCTCTCT 668
Db      181 AlGIYThraLysCyseLeuValSerGIYTrpGIYThraLysSerProGlnValHisPhe 200
      669 GCTGGGACAAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 728
QY      201 ProLysValLeuGlnCyseLeuAsnIleSerValIleuSerGlnLysArgCyseLysPala 220
      729 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGTCTTAAGTCAAGAAAGTGTGAGATGCT 788
QY      221 TyrProArgGlnIleAspAspThraMetPheCyseAlaGIYAspLysArgIYAspAspSer 240
      789 TACCCGACAGACAGATGACATCATATCTTGTGCGCGGTGACAAAGAGATGAGACTCC 848
Db      241 CyseGlnGIYAspSerGIYGIYProValIYCyAsnGIYSerLeuGlnGIYLeuValSer 260
      849 TGCACAGGTGATCTGGGGGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 908
QY      261 TrpGIYAspTrpProCyseAlaArgProAsnArgProGIYValTYThraAsnLeuCyseLys 280
      909 TGGGAGATTACCTTGTGTGCGCGGCGCAACAGACCGGGGTCTTACAGAACTCTGCAAG 968
Db      281 PheThrLysTrpTrpIleGlnGIYThraIleAlaAsnSer 293
      969 TTCAACAAGTGAATCCAGAAACCATCCAGGCCCAACTCC 1007

```

```

RESULT 6
US-09-938-671-52
; Sequence 52, Application US/09938671
; Publication No. US20040002066A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2022P1
; CURRENT APPLICATION NUMBER: US/09/938,671
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/348,457
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: 60/070,567
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,692

```

```

; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,704
; PRIOR FILING DATE: 1998-01-07
; PRIOR APPLICATION NUMBER: 60/070,658
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 52
; LENGTH: 1381
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-938-671-52

Alignment Scores:
Pred. No.: 3,73e-187 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271B-14 (1-293) x US-09-938-671-52 (1-1381)
QY      1 MetAlaThraLysArgProTrpMetCTrValLeuCyseAlaLeuIleThraAlaLeu 20
      129 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTGCTGTGATCACAGCTTCTT 188
QY      21 LeuGIYValThrGIuHIEValLeuAlaAsnAspAspValSerCyAspAspHISProSerAsn 40
      189 CTGGGGGTGCACAGAGCATGTCTTCGCGCAACAATATGTTCTCTGTGACCAACCCCTCTTAC 248
QY      41 ThrValProSerGIYSerAsnGlnAspLeuGIYValaGIYValaGIYValaAspAlaArgSer 60
      249 ACCGTGCTCTTGGAGACAAACAGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAGCTGGAG 308
Db      61 AspAspSerSerSerArgIleIleAsnGIYSerAspCyAspMetHisThrGlnProTrp 80
      309 GATGACAGACAGACCGCATCATCATATGATTCGACATCCGATATGACACCCAGCCCTGG 368
QY      81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTrpCyseGIYValaValLeuValHisPro 100
      369 CAGAGCCCGCTGTGTGCTTAAAGCCCAACAGCTTACGCGGGGGGTGTGTGATCACTCA 428
Db      101 GlnTrpLeuLeuThraAlaAlaHisCyAspArgLysValPheArgValArgLeuGIYHis 120
      429 CAGTGGCTGCTCACGGCGGCGCCACTGCAAGAAAGTTTTCAGAGTCCGTTCTGGGCAC 488
QY      121 TyrSerLeuSerProValTYRGIUSerGIYGIInIleMetPheGlnGIYValIYsSerIle 140
      489 TACTCCCTGTCCACAGTTTATGATCTGGGACAGATGTTTCCAGGGGGTCAAAATCATC 548
Db      141 ProHisProGIYTYRSerHisProGIYHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
      549 CCCACCCCTGCTACTCCACCCCTGGCCACTTAAACCTCATGCTCATCAAACTGAC 608
QY      161 ArgArgIleArgProThraLysAspValArgProIleAsnValSerSerHisCyseProSer 180
      609 AGAAGATTGTGTCCACTTAAAGATGACAGCCCAACAGCTCTCTCATGTTCTCTCT 668
Db      181 AlGIYThraLysCyseLeuValSerGIYTrpGIYThraLysSerProGlnValHisPhe 200
      669 GCTGGACAAAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 728
QY      201 ProLysValLeuGlnCyseLeuAsnIleSerValIleuSerGlnLysArgCyseLysPala 220
      729 CTTAAGGTCTCTCCAGTCTTGAATATCAGCGTCTTAAAGTCAAGAAAGTGTGCGAGATGCT 788
Db      221 TyrProArgGlnIleAspAspThraMetPheCyseAlaGIYAspLysArgIYAspAspSer 240
      789 TACCCGACAGACAGATGACATCATATCTTGTGCGCGGTGACAAAGAGATGAGACTCC 848
QY      241 CyseGlnGIYAspSerGIYGIYProValIYCyAsnGIYSerLeuGlnGIYLeuValSer 260

```

Db 849 TGCAGGGTGAATTCGGGGGCGCTGTGCTCGCATGCTCCCTCGAGGAGACTCGTGTCC 908
QY 261 TTPGLYASPTYrProCySa1aArpProSaNaArgProG1YValYrThraSnLeuCySLys 280
Db 909 TGGGGAGATTACCTTGTGCTGGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAAG 968
QY 281 PheThrLYeTrpLLeGInGluThrLLeGIna1aAsnSer 293
Db 969 TTCACCAAGTGGATCCAGGAACCAATCCAGGCCCACTCC 1007
RESULT 7
US-09-739-907-37
/ Sequence 37, Application US/09739907
/ Patent No. US20010012889A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 36 Human Secreted Proteins
/ FILE REFERENCE: P2022P1
/ CURRENT APPLICATION NUMBER: US/09/739, 907
/ PRIOR FILING DATE: 2000-12-20
/ PRIOR APPLICATION NUMBER: 09/348, 457
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/070, 567
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 692
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 704
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 658
/ PRIOR FILING DATE: 1998-01-07
/ NUMBER OF SEQ ID NOS: 196
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 1516
/ TYPE: DNA
/ FEATURE:
/ ORGANISM: Homo sapiens
/ NAME/KEY: SITE
/ LOCATION: (34)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-739-907-37
Alignment Scores:
Pred. No.: 4,24e-187 Length: 1516
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
Db: 9
US-09-936-271b-14 (1-293) x US-09-739-907-37 (1-1516)
QY 1 MetAlaThra1aArpProTrpMetTrpValLeuCySa1aLeuLLeThra1aLeuLeu 20
Db 254 ATGGCTACAGCAAGACCCCTGGATGTGGGTCTGTGCTGTGTCACAGCTTGTCTT 313
QY 21 LeuG1YValThrG1uH1sValLeu1aAaNaAspValSerCySaSpH1sProSerAa 40
Db 314 CTGGGGGTCAAGAGATGTTCTCGCAACATGATGTTCTCTGACCAACCCCTTAAC 373
QY 41 ThrValProSerG1YSerAaNg1aAspLeuG1YValAG1YG1uAsp1aArGSer 60
Db 374 ACCGTGCTCTGGGAGCAACGAGACTGGGAGCTGGGGCGGGGAGAAAGCGCCGCTCG 433
QY 61 AspAspSerSerSerArgLLe1LeaNg1YSerAaPcySaAspMetH1sThrGInProTrp 80
Db 434 GATGACAGAGACGCGCATCAATCATGATCGACTGCGATATGCAACCCAGCCGTGG 493
QY 81 G1na1aAlaLeuLeuLeuArpProAaNg1nLeuYrCySaG1YValaValLeuValH1sPro 100
Db 494 CAGGCGCGCTGTGCTTAAGGCCCAACAGCTTACTGCGGGGGGATGTGGTGCATCCA 553
QY 101 G1nTrpLeuLeuThra1aAlaH1sCySaArgLYSLysValPheArgValArgLeuG1YH1s 120

Db 554 CAGTGGCTGCTCACCGCGCCCACTGCAGGAAGAAATTTTCAGATCGCTCCGCGCAC 613
QY 121 TyrSerLeuSerProValTYrG1uSerG1YgInG1metPheGInG1YValLYsSerLLe 140
Db 614 TACTCCGTGTCACCGATTATGAATCTGGGAGCAAGATGTTCCAGGGGTCAAAATCCATC 673
QY 141 ProH1sProG1YTYrSerH1sProG1YH1sSerAaNaAspLeuMetLeu1LeYSLeua 160
Db 674 CCCGACCTGGGTACTCCCACTGCGCACTTAAACACCTCATATGCTCATGAACTGAAC 733
QY 161 ArgATg1LeaArgProThrLYeAspValArgProLLeaAnValSerSerH1sCySaProSer 180
Db 734 AGAAGAAATTCGTCACCATTAAGATGTCAAGACCCATCAACGTCCTCTCATATGTCCTCT 793
QY 181 AlaG1YThrLYeCySLeUValSerG1YTrpG1YThrThrLYeSerProG1YValH1sPhe 200
Db 794 GCTGGGACAAAGTGTGTGTGTGTGTGGGGGACCAACCAAGGCCCCCAAGTGCATTC 853
QY 201 ProLYeValLeuGInCySLeuaSn1LeSerValLeuSerG1nLYeArgCySaG1uAspAla 220
Db 854 CCTAAGTCTCTCACTGCTTGAATTCAGCGTCAATCAAGAAAGGTGGAGGATGCT 913
QY 221 TYrProArGgIn1LeAspAspThrMetPheCySa1aG1YAspLYeAlaG1YArgAspSer 240
Db 914 TACCCGAGACAGATGATGACACCAATGTTCTGCGCGGTGCAAGCAGGTAGAGACTCC 973
QY 241 CySaNg1YAspSerG1YG1YProValYalCySaAaNg1YSerLeuGInG1YLeuValSer 260
Db 974 TCCAGGGTGAATTCGGGGGCGCTGTGTGTGTGCAATGAGCTCCCTGCAAGAGATCTGTGCC 1033
QY 261 TTPGLYASPTYrProCySa1aArpProSaNaArgProG1YValYrThraSnLeuCySLys 280
Db 1034 TGGGAGATTACCTTGTGCTGGCGGCCCAACAGACCGGGTGTCTAACAAGACTCTGCAAG 1093
QY 281 PheThrLYeTrpLLeGInGluThrLLeGIna1aAsnSer 293
Db 1094 TTCACCAAGTGGATCCAGGAACCAATCCAGGCCCACTCC 1132
RESULT 8
US-09-938-671-37
/ Sequence 37, Application US/09938671
/ Publication No. US20040002066A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 36 Human Secreted Proteins
/ FILE REFERENCE: P2022P1
/ CURRENT APPLICATION NUMBER: US/09/938, 671
/ PRIOR FILING DATE: 2001-08-27
/ PRIOR APPLICATION NUMBER: 09/348, 457
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: 60/070, 567
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 692
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 704
/ PRIOR FILING DATE: 1998-01-07
/ PRIOR APPLICATION NUMBER: 60/070, 658
/ PRIOR FILING DATE: 1998-01-07
/ NUMBER OF SEQ ID NOS: 196
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 37
/ LENGTH: 1516
/ TYPE: DNA
/ FEATURE:
/ ORGANISM: Homo sapiens
/ NAME/KEY: SITE
/ LOCATION: (34)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-938-671-37
Alignment Scores:
Pred. No.: 4,24e-187 Length: 1516

Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-938-671-37 (1-1516)

QY 1 MetAlaThrAlaArgProTyrMetTyrValLeuCyAlaLeuLeuLeuLeuLeu 20
DB 254 ATGGCTCAGACAGACCCCTGATGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 313
QY 21 LeuGlyValThrGluHisValLeuAlaAspAspValSerCyAspHisProSerAsn 40
DB 314 CTGGGGGTTCACAGACATGTTCTGCGCAACATATGTTCTGTCGACACCCCTCTAAC 373
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSer 60
DB 374 ACCGTGCTCTGTGGAGACACACAGACCTGGGCGCGGGAGAACCGCGGTGCG 433
QY 61 AspAspSerSerSerAlaGlyLeuAlaAsnGlySerAspCyAspMetHisThrGlnProTyr 80
DB 434 GATACAGACAGACCCCATCATCATCAATGATCCGATGCGATATGACACACCCCGGTGG 493
QY 81 GlnAlaAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
DB 494 CAGCGCCGCTGTTGCTTAAGCCCAACAGCTTACTGCGGGGCGGTGTGTCTCTCA 553
QY 101 GlnThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 554 CAGGCGCTGTCTCAGCGCGCGCCCACTGCAGAGAAAGTTTCAAGTCCGCTGCGCAC 613
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValSerTyr 140
DB 614 TACTCCCTGTACACAGTTTATGATCTGGGACAGATGTTCCAGGGGGTCAATTCATC 673
QY 141 ProHisProGlyTyrSerHisAspProGlyHisSerAspAspMetLeuLeuLeuLeu 160
DB 674 CCCCACTGCTGCTCTCACTGCGCCCTGCACTTAACACCTCATGCTCATTAACCTAAC 733
QY 161 ArgArgGlnLeuArgProThrLeuAspValArgProLeuAsnValSerSerHisCyAspProSer 180
DB 734 AGAAGATTCGTCCCACTAAAGATGTCAGAACCTCATGCTCTCATTTGCTCTCT 793
QY 181 AlAGlyThrLeuCyAlaLeuValSerGlyTyrGlyThrThrLeuSerProGlnValHisPhe 200
DB 794 GCTGGGCAAAAGTCTGTGTCTGTCTGTGGGCGACCAACCAAGCCCCCAAGTGCATCTC 853
QY 201 ProLeuValLeuGlnCyAlaLeuAsnHisSerValLeuSerGlnLeuAspValAspAla 220
DB 854 CTTAAGGTCTCTCCAGTCTGTAATATACGCTGTAAGTCAAGAAAGTGCAGAGATGCT 913
QY 221 TyrProArgGlnLeuAspAspThrMetPheCyAlaGlyAspValAspValArgAspSer 240
DB 914 TACCCGAGACAGATATGACACCATTTCTGCGCCCGCTGACAAAGACGATAGATCTC 973
QY 241 CyGlnGlnLeuAspSerGlyGlyProValValCyAlaAsnGlySerLeuGlnGlyValSer 260
DB 974 TGCACAGGTGATTCCTGGGGGCGCTGTGTCTGTGCAATGCTCCCTGCGAGGATCTGTCTC 1033
QY 261 TrpGlyAspTyrProCyAlaArgProAspAspArgProGlyValTyrThrAsnLeuCyAla 280
DB 1034 TGGGAGATTAACCTTGTGCGCGCGCCCAACAGACGGGGTGTCTACAGAACTCTGCAAG 1093
QY 281 PheThrLeuTyrTrpLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 293
DB 1094 TTCACCAAGTGCATCCAGAAACCATCATCAGGCCCAACTCC 1132

RESULT 9
US-09-989-722-308
Sequence 308, Application US/09989722
Patent No. US20020072067A1
GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bockstein, David
APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P27301C63
CURRENT APPLICATION NUMBER: US/09/989,722
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089588
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 4,456-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-989-722-308 (1-1570)

QY 1 MetAlaThrAlaArgProffTpmETrpValLeuCyAlaLeuIleThraAlaLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTATACACCCCTTCTT 413
QY 21 LeuGIyValThrGIuHfIeValLeuAlaAsnAspValIserCyAspHisProSerAen 40
DB 414 CTGGGGGTACAGACATGTTTCTGCAACATGATGTTTCTGTGACCAACCCCTTAC 473
QY 41 ThrValProSerGIySerfAsnGIAspLeuGIyAlaGIyAlaGIyGIuAspAlaArgSer 60
DB 474 ACCGTGCCCTCTGGAGACAACAGGACCTGGGAGCTGGGGCCGGGGAAGACCCGGGTGG 533

QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80
DB 534 GATGACGACGACGACCCCATCATATGATCCAGTCGATATGACACCCGCGGG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
DB 594 CAGGCCGCGCTGTTGCTAAAGGCCCAACAGAGTCTACGCGGGGGGGTGGTGCATCCA 653
QY 101 GlnTTrpLeuLeuTrrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
DB 654 CAGTGGCTGCTCAGCGCGCCCACTGCAGAGAAAGTTTCAGAGTCGCTCGGCGAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerTlle 140
DB 714 TACTCCCTGTGACAGTTTATGAAATCTGGGGAGCAGATGTTCCAGGGGGTCTAAATTCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
DB 774 CCCACCCCTGGCTACTCCACCCCTGGCCACTCTAACCACTCATGCTCAATCAATGAC 833
QY 161 ArgArgGlyLeuArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
DB 834 AGAAGATTCGTCCTCCACTAAAGATGTCAGACCCCATCAAGCTCTCTCATTTGCCCTCT 893
QY 181 AlIGlyThrLysCysLeuValSerGlyTrrGlyThrThrLysSerProGlnValHisPhe 200
DB 894 GCTGGGCAAGAGCTGGTGTGTGCTGGGGGCAACCAAGAGCCCCCAAGTGCATCTC 953
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGlyAspAla 220
DB 954 CTTAAGGTCTCTCCAGTCTTGAATATACGCTGTAAAGTCAAGAAAGTTCGAGAGTCT 1013
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
DB 1014 TACCCGAGACAGATGATGACCATGTCTGCGCCCGTACAAAGCAGGTAGAGATCC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TGCAGGGGTATCTGGGGGGGCGCTGTGTCATGCTGCTCCGCGAGGATCTGTGTC 1133
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
DB 1134 TGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACTCTTGCAAG 1193
QY 281 PheThrLysTrrPileGlnGlnThrIleGlnAlaAsnSer 293
DB 1194 TTCACCAAGTGATCCAGAAACCATCCAGGCCCAATCC 1232

RESULT 10
US-09-989-723-308
Sequence 308, Application US/09989723
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltzen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napiet, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC62
CURRENT APPLICATION NUMBER: US/09/989,723
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088858
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088861
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088876
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/089105
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/089440
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089512
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089514
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089532
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089538
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089598
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089599
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089600
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089653
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089801
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089907
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089908
/ PRIOR FILING DATE: 1998-06-18
/ PRIOR APPLICATION NUMBER: 60/089947
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089948
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/089952
/ PRIOR FILING DATE: 1998-06-19
/ PRIOR APPLICATION NUMBER: 60/090246
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090252
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090254
/ PRIOR FILING DATE: 1998-06-22
/ PRIOR APPLICATION NUMBER: 60/090349
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090355
/ PRIOR FILING DATE: 1998-06-23
/ PRIOR APPLICATION NUMBER: 60/090429
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090431
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090435
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090444
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090445
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090472
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090535
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090540
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090542
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090557
/ PRIOR FILING DATE: 1998-06-24
/ PRIOR APPLICATION NUMBER: 60/090676
/ PRIOR FILING DATE: 1998-06-25

/ PRIOR APPLICATION NUMBER: 60/090678
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090690
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090862
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090863
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/091360
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091478
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091544
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 4,45e-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-936-271b-14 (1-293) x US-09-989-723-308 (1-1570)

QY 1 MetalaThraAlaArgProProTfPmEtrPValLeuCyAlaLeuIeThraAlaLeu 20
Db 354 ATGGCTACAGCAAGACCCCTCGATGTGGGTCTGTGTCTGTATACAGCCTTGGT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
Db 414 CTGGGGGTACAGACGATGTTCTCGCAATGATGTTTCTGTGACCACCCCTCTAAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAlaAspLeu 60
Db 474 ACCGTGCCCTCTGGAGCAACCGAGCTGGAGCTGGAGCCGGGGAAGACGCCGGTGG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80
Db 534 GATGACACGACGACGCCGATCATCATGATCCGACTCGATATGCAACCCAGCCGTGG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 594 CAGGCCGCGCTGTCTTAAGGCCCAACGACTCTACTCGGGGCGGTGTGGTGCATCCA 653
QY 101 GlnTrrPleuLeuThrAlaAlaHisCysArgArgLysValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGCTGTCTACGCGCCGCACTGCAAGAAAGATTTTCAAGTCCGTCTCGACAC 713
QY 121 TyrSerLeuSerProValTyrGlnuSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 714 TACTCCCTGTCAACGATTATGAAATCTGGGACGACGATGTTCCAGGGGTCAAAATCCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160

Db 774 CCCACCCCTGGCTACTCCACCTGGCCACTCTAAGCAGCTCATGCTCAACATGTAAC 833
Qy 161 ArgArgIleArgProThrIleValAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGAAATTCCTCCACTCAAGATGTCAGACCCATCAACGCTCTCTCATTTGCTCTCT 893
Qy 181 AlaGlyThrIleValLeuValSerGlyTyrGlyIleThrIleValSerProGlyIleValHisPhe 200
Db 894 CCTGGGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953
Qy 201 ProIleValLeuGlnCysLeuAsnIleSerValLeuSerGlnIleValArgCysGlyAspAla 220
Db 954 CCTAAGGCTCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1013
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspIleValArgAspSer 240
Db 1014 TACCCGAGACAGATGATGACACCATGTTCTGCCCCGTGCAAGCAGGTAGAGACTCC 1073
Qy 241 CysGlnIleValSerSerGlyIleProValValCysAsnIleSerLeuGlnIleValSer 260
Db 1074 TGCGAGGCTGATTTCTGGGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
Qy 261 TyrIleValProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysAla 280
Db 1134 TGAGGAGATTAACCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1193
Qy 281 PheThrIleTyrPheIleGlnIleThrIleGlnIleAlaIleAsnSer 293
Db 1194 TTCACCAAGTGATCCAGAAACCATCCAGGCCAATCTCC 1232

RESULT 11
US-09-989-279-308
Sequence 308, Application US/09989279
Patent No. US20072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bedstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlesen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2730P1C56
CURRENT APPLICATION NUMBER: US/09/989,279
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17

1 PRIOR APPLICATION NUMBER: 60/089538
2 PRIOR FILING DATE: 1998-06-17
3 PRIOR APPLICATION NUMBER: 60/089598
4 PRIOR FILING DATE: 1998-06-17
5 PRIOR APPLICATION NUMBER: 60/089599
6 PRIOR FILING DATE: 1998-06-17
7 PRIOR APPLICATION NUMBER: 60/089600
8 PRIOR FILING DATE: 1998-06-17
9 PRIOR APPLICATION NUMBER: 60/089653
10 PRIOR FILING DATE: 1998-06-17
11 PRIOR APPLICATION NUMBER: 60/089801
12 PRIOR FILING DATE: 1998-06-18
13 PRIOR APPLICATION NUMBER: 60/089907
14 PRIOR FILING DATE: 1998-06-18
15 PRIOR APPLICATION NUMBER: 60/089908
16 PRIOR FILING DATE: 1998-06-18
17 PRIOR APPLICATION NUMBER: 60/089947
18 PRIOR FILING DATE: 1998-06-19
19 PRIOR APPLICATION NUMBER: 60/089948
20 PRIOR FILING DATE: 1998-06-19
21 PRIOR APPLICATION NUMBER: 60/089952
22 PRIOR FILING DATE: 1998-06-19
23 PRIOR APPLICATION NUMBER: 60/090246
24 PRIOR FILING DATE: 1998-06-22
25 PRIOR APPLICATION NUMBER: 60/090252
26 PRIOR FILING DATE: 1998-06-22
27 PRIOR APPLICATION NUMBER: 60/090254
28 PRIOR FILING DATE: 1998-06-22
29 PRIOR APPLICATION NUMBER: 60/090349
30 PRIOR FILING DATE: 1998-06-23
31 PRIOR APPLICATION NUMBER: 60/090355
32 PRIOR FILING DATE: 1998-06-23
33 PRIOR APPLICATION NUMBER: 60/090429
34 PRIOR FILING DATE: 1998-06-24
35 PRIOR APPLICATION NUMBER: 60/090441
36 PRIOR FILING DATE: 1998-06-24
37 PRIOR APPLICATION NUMBER: 60/090435
38 PRIOR FILING DATE: 1998-06-24
39 PRIOR APPLICATION NUMBER: 60/090444
40 PRIOR FILING DATE: 1998-06-24
41 PRIOR APPLICATION NUMBER: 60/090445
42 PRIOR FILING DATE: 1998-06-24
43 PRIOR APPLICATION NUMBER: 60/090472
44 PRIOR FILING DATE: 1998-06-24
45 PRIOR APPLICATION NUMBER: 60/090535
46 PRIOR FILING DATE: 1998-06-24
47 PRIOR APPLICATION NUMBER: 60/090540
48 PRIOR FILING DATE: 1998-06-24
49 PRIOR APPLICATION NUMBER: 60/090542
50 PRIOR FILING DATE: 1998-06-24
51 PRIOR APPLICATION NUMBER: 60/090557
52 PRIOR FILING DATE: 1998-06-24
53 PRIOR APPLICATION NUMBER: 60/090676
54 PRIOR FILING DATE: 1998-06-25
55 PRIOR APPLICATION NUMBER: 60/090678
56 PRIOR FILING DATE: 1998-06-25
57 PRIOR APPLICATION NUMBER: 60/090690
58 PRIOR FILING DATE: 1998-06-25
59 PRIOR APPLICATION NUMBER: 60/090694
60 PRIOR FILING DATE: 1998-06-25
61 PRIOR APPLICATION NUMBER: 60/090695
62 PRIOR FILING DATE: 1998-06-25
63 PRIOR APPLICATION NUMBER: 60/090696
64 PRIOR FILING DATE: 1998-06-25
65 PRIOR APPLICATION NUMBER: 60/090862
66 PRIOR FILING DATE: 1998-06-26
67 PRIOR APPLICATION NUMBER: 60/090863
68 PRIOR FILING DATE: 1998-06-26
69 PRIOR APPLICATION NUMBER: 60/091360
70 PRIOR FILING DATE: 1998-07-01
71 PRIOR APPLICATION NUMBER: 60/091478
72 PRIOR FILING DATE: 1998-07-02
73 PRIOR APPLICATION NUMBER: 60/091544

1 PRIOR FILING DATE: 1998-07-01
2 PRIOR APPLICATION NUMBER: 60/091519
3 PRIOR FILING DATE: 1998-07-02
4 PRIOR APPLICATION NUMBER: 60/091626
5 PRIOR FILING DATE: 1998-07-02
6 PRIOR APPLICATION NUMBER: 60/091633
7 PRIOR FILING DATE: 1998-07-02
8 PRIOR APPLICATION NUMBER: 60/091978
9 PRIOR FILING DATE: 1998-07-07
10 PRIOR APPLICATION NUMBER: 60/091982
11 PRIOR FILING DATE: 1998-07-07
12 PRIOR APPLICATION NUMBER: 60/092182
13 PRIOR FILING DATE: 1998-07-09

Alignment Scores:

pred. No.:	4 45e-187	length:	1570
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-936-271B-14 (1-293) x US-09-989-279-308 (1-1570)

QY 1 MetalaThraAlaArgProTfPmetTfPValLeuCyAlaLeuIleThraIleLeu 20
354 ATGGCTACGCAAGACCCCTGATGTGGTGTCTGTGCTGATCACAGCTTGGCTT 413

QY 21 LeuGlyValThGlnHisValLeuAlaMetAlaMetValSerCyAspHisProSerL 40
414 CTGGGGGCTCACAGACATGTTCTCCCAACATGATGTTCTTGACCAACCCCTTAC 473

QY 41 ThrValProSerGlySerAlaGlnLeuGlyAlaGlyAlaGlyAlaArgSer 60
474 ACCGTGCTCTTGAGACACACGACCTGAGACTGGGGCCGGGAGAACCCGGTGG 533

QY 61 AspAspSerSerSerArgIleIleAlaGlnIleAlaGlnIleAlaGlnIleProTfP 80
534 GATGACAGCAGCAGCGCATCATGATGATGATGATGATGATGATGATGATGATGAT 593

QY 81 GlnAlaAlaLeuLeuLeuArgProAlaGlnIleLeuIleValLeuValIlePro 100
594 CAGGCCGCGCTGTCTTAAGGCCCAACGCTCTACTGGGGCGGTGTGTGATGATCA 653

QY 101 GlnTfPLeuLeuThraAlaIleHisCyArgGlyValAlaPheArgValArgLeuGlyHis 120
654 CAGTGGCTCTCACGCGCGCCACCTGCGAGAAAGATTTCAGAGTCCGTCTCGGCAC 713

QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnIleMetPheGlnIleValIleSerIle 140
714 TACTCCCTGTACCAAGTTATGATGATGATGATGATGATGATGATGATGATGAT 773

QY 141 ProHisProGlyTyrSerHisProGlyHisSerAlaMetLeuIleLeuIleValLeuAn 160
774 CCCACCTCGGTACTCCACCTGCGCACTTAACGACCTGATGATGATGATGATGAT 833

QY 161 ArgArgIleArgProThrLysAspValArgProIleAlaMetValSerSerHisCyProSer 180
834 AGAAGATTCTGCCACCTTAAGATGATGATGATGATGATGATGATGATGATGAT 893

QY 181 AlaGlyThrLysCyAlaValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
894 GCTGGGACAAAGT 953

QY 201 ProLysValLeuGlnCyAlaValSerValLeuSerGlnLysArgCyGlnValAla 220
954 CTTAAGTCTCTCAGGCTTGAATATGACCGCTGAAGTCAAGAAAAGGTCAGGATGT 1013

QY 221 TyrProArgGlnIleAspAspThrMetPheCyAlaGlyAspValAlaGlyArgAspSer 240
1014 TACCCGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1073

QY 241 CyGlnGlnIleAspSerGlyGlyProValValCyAspGlnIleValLeuValSer 260

Db 1074 TGCAGAGTATTTGGGGGGCTGTGTCGATGGCTCCCGCAGGACTGTGTC 1133
Qy 261 TTPGLApyTPyProCyaaIaXpPbaaMrpProGlyValIYrThraenLeuCyals 280
Db 1134 TGGGAGATTAACCTTGCGCCCGGCCCAAGACGGGTGTCTACAGAACTCTGCAAG 1193
Qy 281 PheThryrTPpIleGIngluThrIleGlnIlaaenber 293
Db 1194 TTCACCAAGTGATTCAGAAACATCCAGGCCAATCC 1232

RESULT 12
US-09-989-727-308
Sequence 308, Application US/09989727
Patent No. US20020072497A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bostein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Bacon, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlesen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tuma, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1G65
CURRENT APPLICATION NUMBER: US/09/989,727
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03

PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089633
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948

; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089952
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/090246
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090252
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090254
; PRIOR FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: 60/090349
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090355
; PRIOR FILING DATE: 1998-06-23
; PRIOR APPLICATION NUMBER: 60/090429
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090431
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090435
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090444
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090445
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090535
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090540
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090542
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090676
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090678
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090690
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090694
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090696
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090862
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/090863
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: 60/091360
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091478
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091544
; PRIOR FILING DATE: 1998-07-01
; PRIOR APPLICATION NUMBER: 60/091519
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091626
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091633
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: 60/091978
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/092182
; PRIOR FILING DATE: 1998-07-09

Alignment Scores:
Pred. No.: 4,45e-187
Score: 1608.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Length: 1570
Matches: 293
Conservative: 0
Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-936-271b-14 (1-293) x US-09-989-727-308 (1-1570)
QY 1 Metalaatharalaaarproprotrpmetrtpvalleucyvalaleuilethralaleu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGGGGCTGTGGCTGTGATCACACCTTGGCTT 413
QY 21 leuglyvalathrghlwhlvalleuallabanaarprovalsercyasphlser 40
DB 414 CTGGGGGTACAGACAGATCTCCCAACAAATGATTTCTGTGACCACTCTTAC 473
QY 41 ThrValProserGlySerAsnGlnAspleuglyalaglyalaglylualaaar 60
DB 474 ACCGTGCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAGAACCCCGGTG 533
QY 61 AapAapSerSerSerArglllellaanglyseraapcyaspmehisthiglnprotr 80
DB 534 GATGACAGCAGACGCCGATCATCATGATCCGATGATATACACCCAGCGGTGG 593
QY 81 GlnlaaialaleuileuabrgproabnglnleuYrCyaglyalValleuValhpro 100
DB 594 CAGGCCGCGCTGTCTTAAGSCCAACAGCTTACTGGGGCCGCTGTGGATGCA 653
QY 101 GlnTrpLeuLeuThralaalahiscybarglyslyvalpheaargvalargleuglyhis 120
DB 654 CAGTGGCTGTCAACGCCGCCCACTGAGGAAGAAATTTTCAAGTCCGTCTGGGCAC 713
QY 121 TyrSerLeuSerProvalTyrglusergylgnglnmetpneuglylvallyserlle 140
DB 714 TACTCCCTGTACACAGTTTATGATGAGGAGCAAGATGTTCCAGGGGGTCAATTCATC 773
QY 141 ProhlsproglYtyrSerhisproglYhisseraanaabmetleuileysleu 160
DB 774 CCCACCTGCTACTCCCACTGSCCACTTAACAGACTATGCTCACTCAACTGAAAC 833
QY 161 ArgArgllearprothrlyasparvalargprolleasnaValserSerhisCyasproser 180
DB 834 AAGAGAAATTCGTCACATTAAGATGACAGCAACCAACGATCTCTTATGTCCTCT 893
QY 181 AlaGlyThrlyscyaleuvalsergyltyrplythrlYserProGlnValhisphe 200
DB 894 GCTGGGACAAAGTGTGGTGTCTGGCTGGGGGAGCAACCAAGAGCCCCAAGTGCATTC 953
QY 201 PolyvalleuGlnCyaleuasnllleservalleuserGlnlyasrCyegluabp1a 220
DB 954 CCTAAGGTCTCTCAATGCTTGAATATCAGCGTCTAAGTCAGAAAAGGTGCGAGATGCT 1013
QY 221 TyrProArgGlnlleasparThrmecPheCyalaGlyasplyalaglyarGaspser 240
DB 1014 TACCCGAGACAGATGATGACACCATGTTCTGGCGCGGAGCAAGAGGTGAGACATCC 1073
QY 241 CygGlnGlyAspSerGlyGlyProvalValCyasanglyserleuglnGlyleuValser 260
DB 1074 TCCCAAGGTGATTCGGGGGGCTGTGGCTGCAATGCTCTCCCTCAGAGGACTCGTGTCC 1133
QY 261 TrpGlyAspTyrrProCyalaalargproanaargproglYvalTYrThrasnleucyls 280
DB 1134 TGGGAGATTACCTTGTGCCCCGAGCAACAGCGGGGTGTACACGAACCTTGCAAG 1193
QY 281 PheThrlyeTrpIleGlnGlnThrIleGlnlaaanser 293
DB 1194 TTCACCAAGTGATCCAGGAACATCCAGGCACTCC 1232
RESULT 13
US-09-989-731-308
; Sequence 308, Application US/09989731
; Patent No. US20020103125A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David

APPLICANT: Deenoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gertlsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Peoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tuma, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202

PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P27301C57
CURRENT FILING DATE: 2001-11-19
CURRENT FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690

/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090694
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090695
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090696
/ PRIOR FILING DATE: 1998-06-25
/ PRIOR APPLICATION NUMBER: 60/090697
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090698
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090699
/ PRIOR FILING DATE: 1998-06-26
/ PRIOR APPLICATION NUMBER: 60/090700
/ PRIOR FILING DATE: 1998-07-01
/ PRIOR APPLICATION NUMBER: 60/091519
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091626
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091633
/ PRIOR FILING DATE: 1998-07-02
/ PRIOR APPLICATION NUMBER: 60/091978
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/091982
/ PRIOR FILING DATE: 1998-07-07
/ PRIOR APPLICATION NUMBER: 60/092182
/ PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.: 4,456-187 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271b-14 (1-293) x US-09-989-732-308 (1-1570)

QY 1 MetAlaThAlAArgProPOTTPMettPvAlLeuCYeAlAleuIethAlaleuLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGAAGTGGGTCTGTGCTGTGATCAGCCCTTGCCT 413
QY 21 LeuGlVAlThGluHIsVAlleuAlaAsnAspVAlSerCYaAspHisProSerAn 40
DB 414 CTGGGGGTCAACAGCATGTTCTCGCAACAATGATGTTCTGTGACCACTCTTAAAC 473
QY 41 ThValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGlnAspAlaArgSer 60
DB 474 ACCGTGCTCTGTGGAGCAACCAAGCACTGGAGGCTGGGGCCGGGGAAGCCGCGTGG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCYaAspMetHisThrGlnProTTP 80
DB 534 GAAGACAGACAGCGCGCATCATCAATGATCGACTGCGATATGCAACCCGCGGTGG 593
QY 81 GlnAlAlaLeuLeuLeuArgProAsnGlnLeuTYrCYeGlyAlaValleuValHisPro 100
DB 594 CAGGCGGCGGTGTCTGCTAAGCCCAACCACTTACTGCGGGCGGTGTGGTCAATCA 653
QY 101 GlnTrpLeuLeuThAlaAlaHisCYaArgGlyAlaValaPheArgValaArgLeuGlyHis 120
DB 654 CAGTGGCTGTCTCAGCGCGCGCACTGCAAGAAAGATTTCAGAGTCCGTCTCGGCAC 713
QY 121 TyrSerLeuSerProValTYrGlnSerGlyGlnGlnMetPheGlnGlnGlyValysSerIle 140
DB 714 TACTCCCTGTCAACCAATTAATGATCTGGGCAACAATGTTCCAGGGGGGTCAATCAATCATC 773
QY 141 ProHisProGlyTYrSerHisProGlyHisSerAsnAspLeuMetLeuIleValysLeuAn 160
DB 774 CCCACCTGTGGCTACTCCCACTGGGCACTTAAGCACTCAATGCTCATCAACATGAAAC 833
QY 161 ArgArgIleArgProThrIlyAspValaArgProIleAsnValSerSerHisCYaProSer 180

DB 834 AAGAAGATTGCTCCACATTAAGATGTCAAGCCCATCAAGTCTCTCATTTGCTCCT 893
QY 181 AlaGlyThrIlyCYeLeuValSerGlyTYrGlyTYrThrIlySerProGlnAlaHisPhe 200
DB 894 GCTGGGACAAAGTGTGGTGTCTGTGGTGGGGAACCAAGGCCCAATGACCTTC 953
QY 201 PolyValleuGlnCYeLeuAsnIleSerValleuSerGlnIlyArgCYeGlnAspAla 220
DB 954 CTTAAGTCTCTCAATGCTTGAATATCAACCGTCTTAAGTCAGAAAAGTGCAGATGCT 1013
QY 221 TYrProArgGlnIleAspAspThrMetPheCYeAlaGlyAspIlyAlaArgAspSer 240
DB 1014 TACCGAGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGAGTAGACCTCC 1073
QY 241 CYeGlnGlyAspSerGlyGlyProValaValCYeAsnGlySerLeuGlnGlyValysSer 260
DB 1074 TCCAGGGGTGATTCGGGGGCTGTGTGTGCAATGCTCTCTGAGGGAGCTGTGTC 1133
QY 261 TTPGlyAspTYrProCYeAlaArgProAsnArgProGlyValaTYrThrAsnLeuCYeHis 280
DB 1134 TGGGAGATTTACCTTGTGCTGCGGCCCAACAGACCGGGGTCTACAGACCTCTGCAAG 1193

RESULT 15

US-09-991-073-308
/ Sequence 308, Application US/09991073
/ Patent No. US20020127576A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi J.
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Eaton, Dan L.
/ APPLICANT: Ferrara, Napoleone
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Kijavini, Ivar J.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ FILE REFERENCE: P2730PLC15
/ CURRENT APPLICATION NUMBER: US/09/991,073
/ PRIOR FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: 60/049787
/ PRIOR FILING DATE: 1997-06-16
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/065186
/ PRIOR FILING DATE: 1997-11-12
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066770
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/075945
/ PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Alignment Scores:

Pred. No.:	4,45e-187	Length:	1570
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-936-271B-14 (1-293) x US-09-991-073-308 (1-1570)

QY 1 MetAlaThrAlaArgProProMetTTPValLeuCyAlaLeuLerThrAlaLeu 20
DB 354 ATGGCTACAGCAAGACCCCTGGATGAGGCTGCTGATCAAGCTTGCTT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 414 CTGGGGGTCAAGAGCATGTTCTGCCAACAATGATGTTCCGTGACCACTTAAAC 473
QY 41 ThrValProSerGlySerSerGlnAspLeuGlyValGlyValGlyGluAspAlaArgSer 60
DB 474 ACCGTGCTCTGGGAGCAACGAGACTGGAGCTGGGGCCGGGAGACGCGCGGTG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
DB 534 GATGACAGACGACCGCATCATCAATGATCCGACTGGATATGCAACCAAGCCGTG 593
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyCyGlyValaValLeuValHisPro 100
DB 594 CAGGCGCGCTGTGTGTAAGGCCCAACAGCTCTATGCGGGGCGGTGTGGTCA 653
QY 101 GlnTrpLeuLeuThrAlaAlaHisCyAspGlyGlyValPheArgValArgLeuGlyHis 120
DB 654 CAGTGCTCTCAAGCGCCGCCCTGACAGAAAGATTTCAGAGTCCGTCTCGGCAC 713
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
DB 714 TACTCCCTGTCAACGATTTATGATCTGGACAGAGATGTTCCAGGGGTCAATCATC 773
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLerLysLeuAsn 160
DB 774 CCCCACTGGGTACTCCACCTGCGCACTCTAACGACCTCATGCTCAAACTGAAC 833
QY 161 ArgArgGlyLeuArgProThrLysAspValArgProIleAsnValSerSerHisCyAspProSer 180
DB 834 AGAAGATTCGTCCCACTAAAGATGCAACCCATCAACGTCTCTCATTTGTCCTCT 893
QY 181 AlaGlyThrLysCyLeuValSerGlyTTPGlyThrThrLysSerProGlnValHisPhe 200
DB 894 GCTGGGACAAAGCTTGGGTGTGGCTGGGGAACAACAAGAGCCCCCAAGTGCATTC 953
QY 201 ProLysValLeuGlnCyLeuAsnLieserValLeuSerGlnLysArgCyGluAspAla 220
DB 954 CTTAAGTCTCTCAAGTCTTGAATATCAAGCTCTAAGTCAAGAAAGTGCAGATGCT 1013
QY 221 TyrProArgGlnLiesAspAspThrMetPheCyAlaGlyAspLysAlaGlyArgAspSer 240
DB 1014 TACCCGAGACAGTAAATGACCAATGTTCTGGCGCGTGACAAAGCAGTAGAGACTCC 1073
QY 241 CysGlnGlyAspSerGlyGlyProValValCyAsnGlySerLeuGlnGlyLeuValSer 260
DB 1074 TGCAGGGATGATTCGGGGGCGCTGTGTCTGCAATGGCTCCCTGCAAGGAACTGTGTC 1133

QY 261 TTPGlyAspTyProCyAlaArgProAsnArgProGlyValTyrThrAsnLeuCyAslys 280
DB 1134 TGGGGAGATTACCTTGTGCCCCGCCCAACAGACCGGGGTCTCACGAACTTGCAAG 1193
QY 281 PheThrLysTTPIleGlnGluThrIleGlnAlaAsnSer 293
DB 1194 TTCACCAAGTGATCCAGGAAACCATCCAGGCAACTCC 1232

Search completed: February 26, 2005, 22:39:46
Job time : 641 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 26, 2005, 12:36:36 : Search time 3685 Seconds
(without alignments)
3026.548 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608
Sequence: 1 MATRRPPMMWVLCALITAL.....VYTLCKFTKIQETIQANS 293

Scoring table:

BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 14239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Filtering first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-Q=/cgn2_1/USPTO.epool/US0936271/tunat_23022005_130127_27273/app_query.fasta_1.455
-DB=EST -QFMT=fastlap -SUFFIX=rcs -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0936271.QCGN_1_1_3437_@tunat_23022005_130127_27273 -NCPU=6 -ICPU=3
-NO MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g981.*
9: gb_g982.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1248.5	77.6	960	2	BE899189
2	1158.5	72.0	1240	3	AK003996
3	1143.5	71.1	1054	2	BE747860
4	1103.5	68.6	997	6	BY704245
5	1066.5	66.3	714	2	BE898407
6	1041	64.7	729	2	BE8842755
7	988.5	61.5	776	4	BG680921
8	899	55.9	663	2	BE742856
9	894	55.6	802	4	BG682309

10	891	55.4	677	2	BE388198
11	787	48.9	677	2	BE898385
12	744	46.3	482	2	AM820357
13	734	45.6	498	7	W73168
14	734	45.6	533	4	BM840509
15	734	45.6	577	4	BM838406
16	734	45.6	590	4	BM840511
17	730	45.4	526	4	BM841697
18	724	45.0	405	4	BM760479
19	722	44.9	544	5	BX520121
20	720	44.8	666	4	BM842155
21	719	44.7	607	7	CV026544
22	713	44.3	953	2	BE745465
23	705	43.8	523	4	BM841293
24	698	43.4	512	4	BM830263
25	690.5	42.9	557	4	BM032286
26	680	42.3	809	4	BG680075
27	671	41.7	765	9	AY400701
28	661	41.1	765	9	AY400702
29	655	40.7	457	1	AI385433
30	650.5	40.5	1295	3	AK009360
31	650.5	40.5	1295	3	AK009720
32	641.5	39.9	783	9	AY410898
33	634	39.4	765	9	AY400703
34	632	39.3	1282	3	BC015551
35	631.5	39.3	1269	3	AK009659
36	627	39.0	783	9	AY410900
37	627	39.0	880	6	CB202840
38	621	38.6	898	6	CB204935
39	620	38.6	945	5	BU151248
40	619	38.5	852	6	CB587168
41	618	38.4	1254	3	AK087426
42	612.5	38.1	754	4	BI653899
43	611.5	38.0	1447	3	AY410899
44	597	37.1	1745	3	AK077406
45	597	37.1	1745	3	AK029477

ALIGNMENTS

RESULT 1
LOCUS BE899189 960 bp mRNA linear EST 29-SEP-2000
DEFINITION 601681989F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3952031 5', mRNA sequence.
ACCESSION BE899189
VERSION BE899189.1 GI:10366425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 960)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Straubeberg, Ph.D.
Email: cgaops-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/PLNL at: image.lnl.gov
Plate: LINC821 row: a column: 24
High quality sequence stop: 744.
Location/Qualifiers
1. 960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3952031"
/tissue_type="adenocarcinoma cell line"

FEATURES

source

```

/lab host="DH10B (phage-resistant)"
/clone lib="NH MGC 9"
/notes="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGACGAG(G). Size selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      4,63e-112      Length:      960
Score:          1248.50        Matches:      242
Percent Similarity: 94.25%      Conservative: 4
Best Local Similarity: 92.72%      Mismatches: 13
Query Match:    77.64%         Indels:      4
DB:             2              Gaps:        1

```

US-09-936-271B-14 (1-293) x BE899189 (1-960)

```

QY      32 AspValSerCyAspAphSPProSerAaThValProSerGlySerHngInAapLeuGly 51
      |||
DB       1 GATGTTTCTGTACCAACCCCTTAACACCGTCCCTGGGAGCAACCAAGCACTGGGA 60
QY      52 AlAGlYAlAGlYgLUaSPAlAaRgSerAaSPSerSerAryIleIleAaNGlySer 71
      |||
DB       61 GCTGGGGCCGGGGAAGAGCCCGGATGGATGACAGCAGCCGATCATCAATGAGATCC 120
QY      72 AspCyAspMethIstThrgInProTIPgInAlAlaLeuLeuAaRgProAaNGInLeu 91
      |||
DB      121 GACTGGGATATGACACACCCAGCCGTGGCAGCCGCGTGTGTAAGGCCCAACGAGCTC 180
QY      92 TyrCyGgIYAlAlValLeuValHISProGInTPLeuLeuThrAlaAlaHISCyAaGly 111
      |||
DB      181 TACTGGGGGGCGGTGTGGTGCATCCACAGTGGCTGTCCAGCGCCGCTCCAGAGAG 240
QY      112 LysValPheArgValArgLeuGlyHISgTySerLeuSerProValTyrGIuSerGIyGIn 131
      |||
DB      241 AAAGTTTTCAGAGTCCGTCTGGGCACACTACTCCCTGACCAAGTTATGAAATCTGGGAG 300
QY      132 GInMePheGInGlyValYLySerIleProHISProGIYTySerHISProGIYHISer 151
      |||
DB      301 CAGATGTTCCAGGGGTC-AAATCCATCCCCACCCGTGGCTACCTCCACCTGGCCACTCT 359
QY      152 AsaAapLeuMeLeuIleLysLeuAaAaGArgIleArgProThrLysAaPValAaGPro 171
      |||
DB      360 AACGACCTCATGTCTCAAACTGAACAGAAAGATTCTCCCACTAAAGATGTCAAGACC 419
QY      172 ILAsnValSerSerHISCySPProSerAlAGlYThrLysCySLeuValSerGIYTPGIY 191
      |||
DB      420 ATCAAGCTCTCTCTCATTTGTCTCTGCTGGACAAAGTCTTGATGTGGCTGGGG 479
QY      192 ThrThLysSerProGInValHISpHeProLysValLeuGInCySLeuAaHISerVal 211
      |||
DB      480 ACAACCAAGGCCCCCAAGTGACCTTCCCTAAGGTCCTCCAGTGCTTGAATATCAAGGT 539
QY      212 LeuSerGIuLysArgCySgIUaSPAlAaTyRProArgInIleAaSPThrMetPheCy 231
      |||
DB      540 CTAAAGTCAGAAAGGTGCGGAGGTCTTACCAGACAGATPAATGACACCAATGTTCTGC 599
QY      232 AlAGlYAspLysAlAGlYArGAsPSPCySgInGlyAspSerGIyGIYProValVal-CY 251
      |||
DB      600 GCCGGTGACAAAGCAGGTAGAGCTCTGCCAGGGGATTTCTGGGGGCGCTGGGTCTTG 659
QY      251 SAaNGlySerLeuGInGlyLeuValSerTrpGIYAspTyRProCySLeuAaRgProAaNG 271
      |||
DB      660 CAATGAGCTCCCGCGGAGACTCGGTGCTCCGGAAGATTTCCCTTGTGC-CGGGCAACAG 718
QY      271 gProGIYValTyRThraNLeuCySlySPheThrLysTrpIleGInGInThrIleGInAl 291
      |||
DB      719 AACGGGTGTCTACAGCAACCTTGCAAG--TCACCAAGTGTGTCGGAACAACATCAAGGC 775

```

```

QY      291 a 291
DB       776 A 776

RESULT 2
AK003996
LOCUS
DEFINITION
AK003996          1240 bp      mRNA          linear      HTC 03-APR-2004
Mus musculus 18-day embryo whole body cDNA, RIKEN full-length
enriched library, clone:1110030019 product:weakly similar to
KALIKREIN 5 PRECURSOR (EC 3.4.21.-) (STRATUM CORNEUM IRITIC
ENZYME) (KALIKREIN-LIKE PROTEIN 2) (KLK-L2) [Homo sapiens], full
insert sequence.
AK003996
AK003996.1 GI:12634990
HTC, CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
Garninci, P. and Hayashizaki, Y.
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
AUTHORS
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
2049374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Yamada, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Yoneda, Y., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer
JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
Nature 409, 685-690 (2001)
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
Nature 420, 563-573 (2002)
MEDLINE
11076861
PUBMED
11076861
REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hirooka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, H., Tagami, M., Tagawa, A., Takaue, F., Tanaka, T.,
Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
Yoshino, M., Yamatsu, M. and Hayashizaki, Y.
TITLE
Direct Submission
JOURNAL
Submitted (10-JUL-2000) Yoshinide Hayashizaki, The Institute of

```


COMMENT

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, Riken Genomic Sciences Center (GSC), Riken Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGCGCGCGCACTCGAGTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3']. cDNA was cleaved with XhoI and SctI. Cloning sites, 5' end: SctI; 3' end: XhoI. Host: SOLR.

FEATURES

SOURCE

Location/Qualifiers

1. 1240
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:1110030019"
/db_xref="taxon:10090"
/clone="1110030019"
/issue_type="whole body"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="18-day embryo"
49. 930
/note="unnamed protein product; putative weakly similar to KALLIKREIN 5 PRECURSOR (EC 3.4.21.-) (STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN 2) (KLK-12) [Homo sapiens] (SWISSPROT|Q9Y3J7, evidence: FASTA, 69.7%ID, 99.6%length, match=879)"
/codon_start=1
/protein_id="BAB23113.1"
/db_xref="GI:12834991"
/translation="MARTGHPMKMAATLITLVLGVSEPLVAGDVSSCDNPGSTERS
GTRNDLSTDSKSGEDTRSDSSRIYVNSDQKQAPWQALILGPNLYCGAVALISQ
WLTLAHCRRKPRIRLGHSMSPVYSSGOMQGISLHPGSHGHNDMLIKM
NRKIRDSHVKPVPRIACDCATBETRCVSWGCTSSSHNNPPKVLQCLNTIVSEBIC
KNSYPGQIDKTFMCAGDEBGRDSCQDSGSPVVCNGKLQGLVSMGDPCCAKQRNRPVY
TNLCFVKWIKDTMNSN"
1218. 1223
/note="putative"
1240
/note="putative"

CDS

ORIGIN

polyA_site

/note="putative"

Alignment Scores:

Pred. No.: 4,79e-103 Length: 1240
Score: 1158.50 Matches: 205
Percent Similarity: 83.67% Conservative: 41
Best Local Similarity: 69.73% Mismatches: 45
Query Match: 72.05% Indels: 3
DB: 3 Gaps: 2

US-09-936-271b-14 (1-293) x AK003996 (1-1240)

QY 1 MetAlaThrAlaArgProPOTPTMettTValLeuCYaLaLeuLeuLeuLeu 20
DB 49 ATGCCGAGGACCGGACACCCCTGGAAATGGGCAATGCTACCCCTGATCACACCTCTGTT 108
QY 21 LeuGlyValThrGluHISValLeuAlaAspAspValSerCysAspHisProSerAsn 40
DB 109 CTGGGGGATCTCAGAGCTCTTCTTCTGGGGATCTTCTCTCTTGTGACACCCCTCTGGA 168
QY 41 ThrValProSerGlySerAsnGlnAspLeuGly-----AlaGlyAlaGlyGluAspAla 58
DB 169 ACCGAACCTCTCTGGGACCAACAGAGACTCAGACGAGATTCAGAGCTCTGGGAGAGACACC 228

QY 59 ArgSerAspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGln 78
DB 229 GGTTC---GATGACGAGCTCTCAATTTGTAAATGGGTCAGACATGCGCAAAAGATGCACAG 285
QY 79 ProTrpGlnAlaAlaLeuLeuLeuAlaArgProAsnGlnLeuTyrcysGlyValAlaVal 98
DB 286 CCATGGCAGGGGCCCTCTGCTTCTGGGGCCCAACAGCTGACTCTGGGGCTGTGCTGATC 345
QY 99 HisProGlnTrpLeuLeuThrAlaAlaHisCysArgIleValPheArgValArgLeu 118
DB 346 AGCCACAGAGGCTGCTCAGACAGACACTGCAGAAAGCCAGATTCACAAATCCGCTG 405
QY 119 GlyHisTrpSerLeuSerProValTyrcysGlyGlnGlnMetPheGlnGlyVal 138
DB 406 GGCACCATTCATGTCACCTGCTTGTATGATCTGGGAGAGATGTTCCAGGGAATCAAA 465
QY 139 SerIleProHisProGlyTyrcSerHisProGlyHisSerAspAspLeuMetLeuIle 158
DB 466 TCCATCCCCCACCCTGGTTACTCCACCTGGGCACATCCATGACTCATCATCAA 525
QY 159 LeuAsnArgArgIleArgProThrTyAspValArgProIleAsnValSerSerHisCys 178
DB 526 ATGAACAGAAATCCGTCAGCTCTCACTCAGTAAGCCCGTCGAAATTGCTGTGACTGT 585
QY 179 ProSerAlaGlyThrTyrcCysLeuValSerGlyTrpGlyThrThrTyrcSerProGlnVal 198
DB 586 GCCACCGAGGAGACTGAGTCAGATGTCGTCTGGGGGCAACAGACACACCAT 645
QY 199 HisPheProValValLeuGlnCysLeuAsnIleSerValLeuSerGlnTyrcAspGlu 218
DB 646 AACTTCCCGAAATCTCTCCAGTCTCCGATATTAAGTGTCTGCTGAGAGAGAGTGA 705
QY 219 AspAlaTyrcProArgGlnIleAspAspThrMetPheCysAlaGlyAspValArgIle 238
DB 706 AACTCTCAACCAAGACAGATAGACAGACCATGTTCTGCCAGATGTGAAGGGCAGG 765
QY 239 AspSerCysGlnIleTyrcSerGlyTyrcProValValCysAsnIleSerLeuGlnIle 258
DB 766 GACTCTCGCCAGGAGTATTCGGAGGTCCTGTGCTGCAATGCAATGCAAGGGCTT 825
QY 259 ValSerTrpGlyAspTyrcProCysAlaArgProAsnArgProGlyValTyrcThrAsnLeu 278
DB 826 GTTCTCTGGGGTATTTCTTCTCTGCTGCTCAGCGGAACAGACAGAGTGTCTACACCACTG 885
QY 279 CysLeuPheThrTyrcTrpIleGlnIleThrIleGlnAlaAsn 292
DB 886 TGTGAGTTCGTTAAGTGAATTAAGACACCATGACTCCAACT 927

RESULT 3
BE747860
LOCUS
DEFINITION
601579720P1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3928372 5',
mRNA sequence.
ACCESSION
BE747860
VERSION
BE747860.1 GI:10161852
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1054)
NIH-MGC <http://mgc.ncl.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov.

ORIGIN (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:

Pred. No.:	2,44e-94	Length:	714
Score:	1066.50	Matches:	214
Percent Similarity:	97.72%	Conservative:	0
Best Local Similarity:	97.72%	Mismatches:	4
Query Match:	66.32%	Indels:	5
DB:	2	Gaps:	0

US-09-936-271B-14 (1-293) x BE898407 (1-714)

```
QY      1 MetAlaThAlaArProProTrpMetTrpValLeuCyAlaLeuIleThAlaLeu 20
        |||
DB      64 ATGGCTACAGCAAGACCCCTGGAGTGGTCTGTGCTCTGATCAAGCTTGCTT 123
QY      21 LeuGlyValThrGluHISValLeu-AlaIleAsnAspValSerCyAspHisProSerAs 40
        |||
DB      124 CTGGGGGTCAAGAGCATGTCTCGGCCAATGATGTTTCTGTGACCACTCTTA 183
QY      40 nThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSe 60
        |||
DB      184 CACCGTCCCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAAAGACGCTGGCTC 243
QY      60 rAspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTr 80
        |||
DB      244 GGATGACACAGAGAGCCGATCATCAATGATTCGATCGATATGCACACCCAGCCGTC 303
QY      80 pGlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcGlyAlaValLeuValHisPr 100
        |||
DB      304 GGAGGCGCGCTGTCTGTAAGGCCCAACGAGCTCTAGCGGGCGGTGTGGTGATCC 363
QY      100 oGlnTrpLeuLeuThAlaAlaHisCyAsnGlyValValPheArgValArgLeuGlyHi 120
        |||
DB      364 ACGAGTGGCTCAGCGGCCGACCTGCAGAGAAAGTTTTCAGAGTCCGCTCGGCCA 423
QY      120 sTyrsSerLeuSerProValTyrcGluSerGlyGlnGlnMetPheGlnGlyValIlySerI 140
        |||
DB      424 CTACTCCCTGTACACGATTATGAACTGGGACGAGATGTTCCAGGGGCTC-AAATCCAT 482
QY      140 eProHisProGlyTyrsSerHisProGlyHisSerAsnAspLeuMetLeuIleIlyLeuAs 160
        |||
DB      483 CCCCCACCTGGAGTACTCCACCTGG-CACCTTAACGACCTCATCTCATCAAACTGAA 541
QY      160 nArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisGlySerPro 180
        |||
DB      542 CAGAGAAATTCGTCCTCAAGATGTCAGACCCATCAACGTCCTCTCATTTGTCCTC 601
QY      180 rAlaGlyThrIlyCyAlaLeuValSerGlyTrpGlyThrThrIlySerProGlnValHisPh 200
        |||
DB      602 TGTGTGGAGCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 659
QY      200 eProIlyValLeuGlnCyAlaLeuAsnIleSerValLeuSerGlnIlyArgCyGlu 218
        |||
DB      660 CCTTAAGGTCTCCAGTGTCTGATATTCAGCGGTGAAGTCAGAAAGGTGCGAG 714
RESULT 6
BM842755 729 bp mRNA linear EST 06-MAR-2002
LOCUS K-EST0120216 S12SNU216 Homo sapiens cDNA clone S12SNU216-66-210 5',
DEFINITION mRNA sequence.
ACCESSION BM842755
VERSION BM842755.1 GI:19199164
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 729)
Kim,M.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
```

TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsun@mail.kr.ibm.re.kr
Plate: 66 row: D column: 10
High quality sequence stop: 729.
Location/Qualifiers

FEATURES

source

1..729

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S12SNU216-66-D10"
/sex="F"
/tissue_type="Lymph node"
/cell_type="Epithelial"
/cell_line="SNJ-216"
/lab_host="Top10F"
/clone_lib="S12SNU216"
/note="Organ: Stomach; Vector: pcNS; Site: 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tobacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of 74 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
```

ORIGIN

Alignment Scores:

Pred. No.:	8.1e-92	Length:	729
Score:	1041.00	Matches:	191
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	64.74%	Indels:	0
DB:	4	Gaps:	0

US-09-936-271B-14 (1-293) x BM842755 (1-729)

```
QY      1 MetAlaThAlaArProProTrpMetTrpValLeuCyAlaLeuIleThAlaLeu 20
        |||
DB      156 ATGGCTACAGCAAGACCCCTGGAGTGGTCTGTGCTCTGATCAAGCTTGCTT 215
QY      21 LeuGlyValThrGluHISValLeuAlaIleAsnAspValSerCyAspHisProSerAsn 40
        |||
DB      216 CTGGGGGTCAAGAGCATGTCTCGGCCAATGATGTTTCTGTGACCACTCTTAAC 275
QY      41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyValAspAlaArgSer 60
        |||
DB      276 ACCGTGCTCTGGAGCAACGAGACCTGGAGCTGGGGCCGGGAAAGACCCCGGTG 335
QY      61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTrp 80
        |||
DB      336 GATGACAGAGAGAGCCGATCATCAATGATCGACTGATATGCACACCCAGCCGCTG 395
QY      81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcGlyAlaValLeuValHisPro 100
        |||
DB      396 CAGGCGCGCTTGTCTAAGGCCCAACGCTCTACTGGGGGCGGTGTGTGTGATCCA 455
QY      101 GlnTrpLeuLeuThAlaAlaHisCyAsnGlyValValPheArgValArgLeuGlyHis 120
        |||
```

Db 456 CAGTGGCTGCTCAGCGGCCGCTGAGCAAGAAATTTCAGATCCGCTCGGCGAC 515
 QY 121 TysTerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
 Db 516 TACTCCCTGTCACAGAGTTTGAATCTGGGCGACAGATGTTCCAGGGGGTCAATTCATC 575
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLeuAsn 160
 Db 576 CCCCACTGCTGCTCTCCACCTGCGCACTCTTAACACTCATGCTCATCAACTGAC 635
 QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
 Db 636 AGAAGAATTCGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATGTGCTCT 695
 QY 181 AlaGlyThrIleCysLeuValSerGlyTyrGly 191
 Db 696 GCTGGGCAAGAGTCTGTCTGCTGCTGGG 728
 RESULT 7
 LOCUS BG680921 776 bp mRNA linear EST 01-MAY-2001
 DEFINITION 602648718F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4753469 5',
 mRNA sequence.
 ACCESSION BG680921
 VERSION BG680921.1 GI:13912305
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 776)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: James Cleaver, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM10613 row: C column: 06
 High quality sequence stop: 770.
 FEATURES
 source
 Location/Qualifiers
 1..776
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4753469"
 /tissue_type="squamous cell carcinoma"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NCI CGAP Skn4"
 /notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: Not;
 Site 2: Sali; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.5kb. Library constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."
 ORIGIN
 Alignment Scores:
 Prod. No.: 13e-86 Length: 776
 Score: 988.50 Matches: 197
 Percent Similarity: 93.87% Conservative: 2
 Best Local Similarity: 92.92% Mismatches: 7
 Query Match: 61.47% Indels: 6
 DB: 4 Gaps: 1
 US-09-936-271b-14 (1-293) x BG680921 (1-776)
 QY 1 MetAlaThrAlaArgProPrtPrtMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
 Db 151 ATGGCTACAGACAGACCCCTGGATGTGGGTGCTGTGCTGTGATCAGACGCTTCTT 210

QY 21 LeuGlyValThrGlnHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
 Db 211 CTGGGGGTCAAGAGCATGTTCTCGCAACAAAGATGTTCTGTGACACCCCTCTAAC 270
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyValaGlyValaGlyValaArgSer 60
 Db 271 ACCGTGCTCTCTGGAGCAACAGACCTGGAGCTGGGGCCGGGAAGACCCGGCTG 330
 QY 61 AspAspSerSerSerArgIleIleGlnGlySerAspCysAlaPheHisThrGlnProTrp 80
 Db 331 GATGACAGCAGACCGCCATCATCAATGATCGAGTGCATATGACACCCGCTGG 390
 QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCysGlyValaValLeuValHisPro 100
 Db 391 CAGGCGCGGTGTGTGCTAAAGCCCAACACACTTACTTCCGGCGGTGTGTGATCA 450
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgGlyValValPheArgValaArgLeuGlyHis 120
 Db 451 CAGTGGCTGCTCAAGCGCCGCTGAGGAAAGATTTTCAGAGTCCGCTCGGCGAC 510
 QY 121 TysTerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValIleSerIle 140
 Db 511 TACTCCCTGTCACAGATTATGAATCTGGGCGACAGATGTTCCAGGGGGTCAATTCATC 570
 QY 141 ProHisProGlyTyrTerHis-ProGly-HisSerAsnAsp-LeuMetLeuIleLeu 159
 Db 571 CCCCACTGCTGCTACTCCACGCTGAGCACCTTAACGAGCTCATGCTCATCAACTG 630
 QY 160 -AaaArgArgIleArgProThrIleAspVal-ArgProIleAsnValSerSerHisCysP 179
 Db 631 AAACAGAGATTCGTCCCACTAAAGATGTCAGAACCCATCAACGTCTCTCATATGTC 690
 QY 179 roSerAlaGlyThrIleCysLeuValSerGlyTyrGlyThrIleCysSerProGlnValH 199
 Db 691 CCTCTGCTGGAGCAAGAGTCTGTGTCTGCTGGGCGACAGACCAAGACCCCAAGTGC 750
 QY 199 IaspProIysValLeuGlnCysLeu 207
 Db 751 ACTTCGCTAGG--TCTCAGTCTTG 773
 RESULT 8
 LOCUS BE742856 663 bp mRNA linear EST 15-SEP-2000
 DEFINITION 601574513F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3835490 5',
 mRNA sequence.
 ACCESSION BE742856
 VERSION BE742856.1 GI:10156848
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 663)
 NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
 plate: LLM517 row: J column: 03
 High quality sequence stop: 657.
 FEATURES
 source
 Location/Qualifiers
 1..663
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3835490"

/cissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH-MGC 9"
/notes="Organ: ovary; Vector: pORF7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dt priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCAAGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:

Pred. No.:	6,6e-78	Length:	663
Score:	899.00	Matches:	193
Percent Similarity:	93.72%	Conservative:	1
Best Local Similarity:	93.24%	Mismatches:	10
Query Match:	55.91%	Indels:	8
DB:	2	Gaps:	1

US-09-936-271B-14 (1-293) x BG682309 (1-663)

QY	1	MetAlaThraAlaArgProPOTPMeTTPVAlLeuCyAlaLeuIleThraAlaLeu	20
DB	54	ATGGCTACAGACAGACCCCTGGATGGTCTCTGCTCTGATCACAGCTTGCTT	113
QY	21	LeuGlyValThraGluHisValLeuAlaAsnAspValSerCysAspHisProSerAn	40
DB	114	CTGGGGGTACACAGATGTTCTGGCCAAAGATGTTCTGTATACACCCCTCTTAC	173
QY	41	ThraValProSerGlySerAsnGlnAspLeuGlyAlaGlyValAlaGluAspValSer	60
DB	174	ACCGTCCCTCTGGAGCAACCGACCTGGAGCTGGGGCCGGGAACGCCCGCTG	233
QY	61	AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTP	80
DB	234	GATGACAGACAGACCGCATCATGATGATCGACTGCGATATGACACCCAGCGTGG	293
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyCysGlyValAlaLeuValHisPro	100
DB	294	CAGGCGCGCTGTGTCTAAAGCCCAACAGCTTACTGCGGGGGGTGGTGGTATCCA	353
QY	101	GlnTrpLeuLeuThraAlaHisCysArgGlyValValPheArgValArgLeuGlyHis	120
DB	354	CAGTGGCTGCTCAGCGCCGC-CACTGCAAGAAAGATTTCAGATCCCTCTCGGCCAC	412
QY	121	TyrBerLeuSerProValIlyrGlySerGlyGlnGlnMetPheGlnGlyValIlySerIle	140
DB	413	TACTCCCTGTCACACATTTATGATCTGGGCAAGATGTTCCAGGGGTC-AAATCCATC	471
QY	141	ProHisProGlyTyrSerHisPProGlyHisSerAspAspLeuMetLeuIlySerLeuAs	160
DB	472	CCCCACCTGGATATCCCA--CCTGGCCACTTAAAGACTCAATGCTCACTCAAACTTGA	528
QY	160	ArgArgGlyLeuArgProThraSerAspValArgProIleAsnValSerSerHisCysProse	180
DB	529	CAGAAATTCGTCCTCACTAAAGATGTCAGACATC-AACGTCTCTCTCATATTGATCCTC	587
QY	180	ValGlyThrIlyCysLeuValSerGlyTyrGlyThrThrIlySerProGlnValHisP	200
DB	588	TGCTGGAGCAAAAGT	646
QY	200	heProIlyValLeuGln 205	
DB	647	C-CTTAAGTCTTGTAG 662	

RESULT 9
BG682309 802 bp mRNA linear EST 01-MAY-2001
LOCUS 602629626f1 NC1_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4754263 5',
DEFINITION mRNA sequence.
ACCESSION BG682309

VERSION BG682309.1 GI:13913706
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 (bases 1 to 802)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LAM10615 row: d column: 08
High quality sequence stop: 802.

FEATURES

source

1..802
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4754263"
/cissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Alignment Scores:

Pred. No.:	2.69e-77	Length:	802
Score:	894.00	Matches:	161
Percent Similarity:	99.38%	Conservative:	0
Best Local Similarity:	99.38%	Mismatches:	1
Query Match:	55.60%	Indels:	0
DB:	4	Gaps:	0

US-09-936-271B-14 (1-293) x BG682309 (1-802)

QY	132	GlnMetPheGlnGlyValIlySerIleProHisProGlyTyrSerHisProGlyHisSer	151
DB	3	CAGATGTTCCAGGGGTCAATCCATCCCACTGGCTACTCCACCTGGCCACTCT	62
QY	152	AsnAspLeuMetLeuIleIlySerLeuAsnArgArgProThraSerAspValArgPro	171
DB	63	AACGACCTCATGCTCATCAACTGACAGAAAGATTCCTCCACTTAAGATGTCAAGCC	122
QY	172	IleAsnValSerSerHisCysProSerAlaGlyThrIlyCysValLeuValSerGlyTyrGly	191
DB	123	ATCAAGTCTCTCTCATATGTCCTCTGTGGAGCAAAAGTGTGTGTGTGTGTGTGTGTGT	182
QY	192	ThrThrIlySerProGlnValHisPheProIlyValLeuGlnCysLeuAsnIlySerVal	211
DB	183	ACAACCAAGAGCCCCCAAGTGCATCTCCCTTAAGGTCCTCCAGTGTGTGAATATCAGCCGG	242
QY	212	LeuSerGlnIlyArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPheCys	231
DB	243	CTTAAGTCAGAAAGGCGAGAGATGTTACCGAGCAATAGATGACACCATTTCTCTG	302
QY	232	AlaGlyAspIlyAlaGlyArgAspSerCysGlnGlyAspSerGlyIlyProValValCys	251
DB	303	GCCTGTGACAAAGCAGGTAGAGATCTCTCCAGGGGATTCGGGGGGCTGTGTGTGTGT	362
QY	252	AsnGlySerLeuGlnGlyLeuValSerTyrGlyAspTyrProCysAlaArgProAsnArg	271
DB	363	AATGGCTCCCTCAGGAGCTCGTGTCTGTGGGAGATTACCTTGTGTGTGTGTGTGTGTGT	422

QY 272 ProGlyValThrAspLeuCyLeuPheThrIleGlnIleGlnAla 291
DB 423 CCGGGTCTTACAGCACTCTGCAAGTTCACCAAGTGATCCAGAAACATCCAGGCC 482
QY 292 AsnSer 293
DB 483 AAGTCC 488
RESULT 10
BE388198 677 bp mRNA linear EST 21-JUL-2000
LOCUS 60128704F1 NIH_MGC_44 Homo sapiens cDNA clone IMAGE:3621682 5',
DEFINITION mRNA sequence.
ACCESSION BE388198.1 GI:9333563
VERSION BE388198.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straube, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LICM295 row: m column: 11
High quality sequence stop: 618.
Location/Qualifiers
1. 677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3621682"
/tissue_type="endometrium, adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_44"
/note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
ORIGIN
Alignment Scores:
Pread. No.: 4.17e-77 Length: 677
Score: 891.00 Matches: 176
Percent Similarity: 92.15% Conservative: 0
Best Local Similarity: 92.15% Mismatches: 10
Query Match: 55.41% Indels: 7
DB: 2 Gaps: 1
US-09-936-271b-14 (1-293) x BE388198 (1-677)
QY 1 MetaLathrAlaArgProPOTripMetTyrValLeuCyAlaLeuIleThrAlaLeu 20
DB 101 ATGGCTACAGCAAGCCCTCGATGTGGTCTCTGTCTGATCAGTCCCTTCTT 160
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCyAspHisProSerAsn 40
DB 161 CTGGGGGTCAAGAGCATGTTCTGGCAACATGATGTTCTTCTGACCAACCTCTTAC 220
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60

DB 221 ACCGTGCTCTTGGAGCAACAGACTGGAGCTGGGCGGGGAAGAGCCGGTCC 280
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCyAspMetHisThrGlnProTyr 80
DB 281 GATGACAGCAGACCGGCATCATCATGATGATCCGACTGCGATATGACACCGCCGTGG 340
QY 81 GlnAlaAlaLeuLeuLeuAspProAsnGlnLeuTyrCysGlyAlaValIleLeuAlaHisPro 100
DB 341 CAGGCCGCGCTGTGCTTAAAGCCCAACACACTTACTGCGGGCGGTGGTGCATCA 400
QY 101 GlnThrLeuLeuThrAlaAlaHisCysArgLeuLeuValPheArgValArgLeuGlyHis 120
DB 401 CAGTGGCTGCTCAAGCGCGCCCACTGAGAGAAAGATTTCAGAGTCCGTCTCGGCAC 460
QY 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValSerSerIle 140
DB 461 TACTCCCTGCACAGATTATGATCTGGGACAGATGTTCCAGGGGGTCAATTCCAT 520
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleValAsn 160
DB 521 CCCCACCTGG-TACTCCACCTGGCCACTGACCTTACGACCTGCTCATCAACTAAC 579
QY 161 ArgArgIleArgProThrIleAspValArgProIleAsnValSerSerHisCysProSer 180
DB 580 AGAAGAAATTCCTCCATTAAAGATGTCAGACCATCAACCTCTCTATGATGCTCTG 638
QY 181 AlaGlyThrIleCysLeuValSerGlyTyrGly 191
DB 639 CTGGGA-----AAAGCTGGGT 656
RESULT 11
BE898385 677 bp mRNA linear EST 29-SEP-2000
LOCUS 601681219F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3951365 5',
DEFINITION mRNA sequence.
ACCESSION BE898385.1 GI:10364809
VERSION BE898385.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straube, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LICM619 row: f column: 06
High quality sequence stop: 677.
Location/Qualifiers
1. 677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3951365"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH_MGC_9"
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGCAAGG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit

ORIGIN (Stratagene) and Superscript II RT (Life Technologies)."

Alignment Scores:

Pred. No.:	7,116-67	Length:	677
Score:	787.00	Matches:	143
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48.94%	Indels:	0
DB:	2	Gaps:	0

US-09-936-271b-14 (1-293) x BE898385 (1-677)

```
QY 151 SerAsnAspLeuMetLeuIleValLeuAsnArgAlaArgProThrIleAspValArg 170
    |||||
DB 3 TCCTAAGACCTCTATGCTCATCAACAGAGAGATTGCTCCCTAAAGATGTCAGA 62
    |||||
QY 171 ProIleAsnValSerSerHisCysProSerAlaGlyThrIleCysLeuValSerGlyTTP 190
    |||||
DB 63 CCCATCAAGCTCTCTCTCATTTGCTGCTGGGACAAAGTCTTGCTGGCTGG 122
    |||||
QY 191 GlyThrThrIleSerProGlnValHisPheProIleValLeuGlnCysLeuAsnIleSer 210
    |||||
DB 123 GGGACACCAAGAGCCGCCCAAGTGCACCTTCCCAAGGTCCTCAAGTGAATATCAGC 182
    |||||
QY 211 ValLeuSerGlnIleValArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPhe 230
    |||||
DB 183 GTGCTAAGTCAAGAAAGGTCCGAGATGCTTACCCAGACAGATGATGACACCATGTTCC 242
    |||||
QY 231 CysAlaGlyAspIleValaGlyArgAspSerCysGlnIleValAspSerGlyIleProValIle 250
    |||||
DB 243 TGGCCCGGTGACAAACAGAGTAGAGACTCCTGCGCAGGGTGATTTCTGGGGGCTGTGGTGC 302
    |||||
QY 251 CysAsnGlySerLeuGlnIleValSerTgplValAspTyrProCysAlaArgProAsn 270
    |||||
DB 303 TGGCAATGGCTCCTGACGGGACTGCTGCTGGGGAGATTAACCTTGTGCCGGCCACAC 362
    |||||
QY 271 ArgProGlyValTyrThrAsnLeuCysIlePheThrIleGlnIleGlnIleGln 290
    |||||
DB 363 AGACCGGGGTCTCAACAGAACCTCTGCAAGTTACCAAGTGCATCCAGAAACATCCAG 422
    |||||
QY 291 AlaAsnSer 293
    |||||
DB 423 GCCAACTCC 431
    |||||
```

RESULT 12
AM820357/c 482 bp mRNA linear EST 17-MAY-2000
LOCUS AM820357
DEFINITION QV2-ST0296-100400-130-d07 ST0296 Homo sapiens cDNA, mRNA sequence.
ACCESSION AM820357
VERSION AM820357.1 GI:7913351
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 482)
Dias Nereo, E., Garcia Correa, R., Verjovsky-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, R.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL MEDLINE
PUBMED
COMMENT

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=QV2-ST0296-100
400-130-d07&t3=2000-04-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 62
High quality sequence stop: 482.
Location/Qualifiers
1..482
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0296"
/note="Organ: stomach; Vector: puc18; Site 1: Smal;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESSES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.:	7,416-63	Length:	482
Score:	744.00	Matches:	139
Percent Similarity:	97.90%	Conservative:	1
Best Local Similarity:	97.20%	Mismatches:	2
Query Match:	46.27%	Indels:	1
DB:	2	Gaps:	0

US-09-936-271b-14 (1-293) x AM820357 (1-482)

```
QY 152 AsnAspLeuMetLeuIleValLeuAsnArgAlaArgProThrIleAspValArgPro 171
    |||||
DB 465 AACGACCTCATGCTCATCAACAGAGAGATTGCTCCCACTAAAGATGTCAGACCC 406
    |||||
QY 172 IleAsnValSerSerHisCysProSerAlaGlyThrIleCysLeuValSerGlyTTPgIy 191
    |||||
DB 405 ATCAAGCTCTCTCTCATTTGCTGCTGCGGACAAAGTGTGCTGTGGCTGGGG 346
    |||||
QY 192 ThrThrIleSerProGlnValHisPheProIleValLeuGlnCysLeuAsnIleSerVal 211
    |||||
DB 345 ACACCAAGAGCCGCCCAAGTGCATCTCCCTTAAGTCTTCAAGTGTGATTCACGCTG 286
    |||||
QY 212 LeuSerGlnIleValArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPheCys 231
    |||||
DB 285 CTAAAGTCAAGAAAGGTGCGAGAGATGCTTACCCGAGACAGATGATGACACCATGTTCTCG 226
    |||||
QY 232 AlaGlyAspIleValaGlyArgAspSerCysGlnIleValAspSerGlyIleProValValCys 251
    |||||
DB 225 GCCGGTGAACAAAGCGTGAAGAGATCTTCCACAGGGATTCGGGGGCTGTGCTGCTGC 166
    |||||
QY 252 AangIleSerLeuGlnIleValSerTgplValAspTyrProCysAlaArgProAsnArg 271
    |||||
DB 165 AATGGCTCTCTGACAGGACTGCTGCTGGGAGATTACCTTGTGCCGGCCCAACAGA 106
    |||||
QY 272 ProGlyValTyrThrAsnLeuCysIlePheThrIleGlnIleGlnIleGlnIle 291
    |||||
DB 105 CTGGGGTGTACACGAACTCCGCGAGTTCACCAAGTGAATCCAGGAACATCCAGGC 46
    |||||
QY 291 AlaAsnSer 293
    |||||
DB 45 CAATCTCC 39
    |||||
```

RESULT 13

LOCUS W73168 498 bp mRNA linear EST 16-OCT-1996
DEFINITION z55se11.r1 Soares_fetal_heart_NbH19w Homo sapiens cDNA clone
IMAGE:344588 5' similar to SW:TR2Y_CNFPA P06872 TRYPSINOGN,

Alignment Scores:

Pred. No.: 8,256-62 Length: 533
 Score: 734.00 Matches: 132
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.65% Indels: 0
 DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x BM840509 (1-533)

QY 162 Arg11LeaRgPProThrLyAspValArgPProIleAsnValSerSerHisCysProSerAla 181
 |||||
 DB 3 AGAATTGCTCCCACTAAAGATGCAGACCCATCAAGCTCCCTCATTTGCTCTGCT 62
 QY 182 G1YThrLysCysLeuValSerG1YTPG1YThrThyLysSerProGlnValHisPhePro 201
 |||||
 DB 63 GGGACAAAGTCTGGTGGTGGTGGGAGCAACCAAGAGCCCAAGTGCATTCCT 122
 QY 202 LysValLeuGlnCysLeuAsnLysSerValLeuSerGlnLysArgCysGluAspAlaTy 221
 |||||
 DB 123 AAGGTCTCCAGTCTTGAATATCAAGCTGCTAAGTCAGAAAGTGCAGAGATGCTTAC 182
 QY 222 ProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSerCys 241
 |||||
 DB 183 CCGAGACAGATGATGATACACCATGTTCTGCGCGGTGACAAACAGTGAAGACTCTGC 242
 QY 242 G1NG1YAspSerG1YG1YProValValCysAsnG1YSerLeuGlnG1YLeuValSerTP 261
 |||||
 DB 243 CAGGGGATGCTGGGGGGGCTGGTGTGCAATGGCTCCCTGCAAGGACTCGTCTGG 302
 QY 262 G1YAspTyProCysAlaArgProAsnArgProG1YValTyThrAsnLeuCysLysPhe 281
 |||||
 DB 303 GGAAGATTACCTTGTGCTCCCGGCCCAACAGACGGGTGTACAGAACCTTGCAAGTTC 362
 QY 282 ThrLysTrpIleGlnGluThrIleGlnIleAsnSer 293
 |||||
 DB 363 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCC 398
 RESULT 15
 BM838406 577 bp mRNA linear EST 06-MAR-2002
 LOCUS K-EST0114765 S12SNU216 Homo sapiens cDNA clone S12SNU216-36-A03 5',
 DEFINITION mRNA sequence.
 ACCESSION BM838406
 VERSION BM838406.1 GI:19194615
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 REFERENCE Bukacynska; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 577)
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
 Oh,K.J., Cheong,U.E., Sohn,H.Y., Kim,U.M., Park,H.S., Kim,S. and
 Kim,Y.S.
 21C Frontier Korean EST Project 2001
 JOURNAL Unpublished (2002)
 COMMENT Contact: Kim YS
 Genome Research Center
 Korea Research Institute of Bioscience & Biotechnology
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
 Tel: +82-42-860-4470
 Fax: +82-42-860-4409
 Email: yongsung@mail.kribb.re.kr
 Plate: 36 row: A column: 03
 High quality sequence stop: 577.
 Location/Qualifiers
 1..577
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="S12SNU216-36-A03"
 /sex="F"

/cissue_type="lymph node"
 /cell_type="Epithelial"
 /cell_line="SNU-216"
 /lab_host="Top10P"
 /clone_id="S12SNU216"
 /note="Organ: Stomach; Vector: pCNS; Site 1: EcoRI;
 Site 2: NotI; The poly (A) + RNA was dephosphorylated with
 bacterial alkaline phosphatase (BAP) and then deprotected
 with tobacco acid pyrophosphatase (TAP). The deprotected
 intact mRNA was ligated with DNA-RNA linker including EcoR
 I site by treatment of T4 RNA ligase and the first strand
 cDNA was synthesized from oligo dt-selected mRNA by
 priming with dt-tailed vector. The dt-tailed vector was
 adjusted to have about 60nt. The cDNA vector was
 circularized with E. coli DNA ligase after digestion of
 EcoRI which site is also included in vector. An RNA strand
 converted to a DNA strand by Okayama-Berg method. The
 obtained cDNA vectors were used for transformation of
 competent cells E. coli Top10P by electroporation method.
 The cDNA libraries constructed by this method are
 full-length enriched cDNA library."

ORIGIN

Alignment Scores:
 Pred. No.: 9,256-62 Length: 577
 Score: 734.00 Matches: 132
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.65% Indels: 0
 DB: 4 Gaps: 0

US-09-936-271B-14 (1-293) x BM838406 (1-577)

QY 162 Arg11LeaRgPProThrLyAspValArgPProIleAsnValSerSerHisCysProSerAla 181
 |||||
 DB 3 AGAATTGCTCCCACTAAAGATGCAGACCCATCAAGCTCCCTCATTTGCTCTGCT 62
 QY 182 G1YThrLysCysLeuValSerG1YTPG1YThrThyLysSerProGlnValHisPhePro 201
 |||||
 DB 63 GGGACAAAGTCTGGTGGTGGTGGGAGCAACCAAGAGCCCAAGTGCATTCCT 122
 QY 202 LysValLeuGlnCysLeuAsnLysSerValLeuSerGlnLysArgCysGluAspAlaTy 221
 |||||
 DB 123 AAGGTCTCCAGTCTTGAATATCAAGCTGCTAAGTCAGAAAGTGCAGAGATGCTTAC 182
 QY 222 ProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSerCys 241
 |||||
 DB 183 CCGAGACAGATGATGATACACCATGTTCTGCGCGGTGACAAACAGTGAAGACTCTGC 242
 QY 242 G1NG1YAspSerG1YG1YProValValCysAsnG1YSerLeuGlnG1YLeuValSerTP 261
 |||||
 DB 243 CAGGGTGTCTGGGGGGGCTGGTGTGCAATGGCTCCCTCAGGGACTCGTCTGG 302
 QY 262 G1YAspTyProCysAlaArgProAsnArgProG1YValTyThrAsnLeuCysLysPhe 281
 |||||
 DB 303 GGAAGATTACCTTGTGCTCCCGGCCCAACAGACGGGTGTACAGAACCTTGCAAGTTC 362
 QY 282 ThrLysTrpIleGlnGluThrIleGlnIleAsnSer 293
 |||||
 DB 363 ACCAAGTGATCCAGAAACCATCCAGGCCAACTCC 398

Search completed: February 26, 2005, 22:29:06
 Job time : 3693 secs